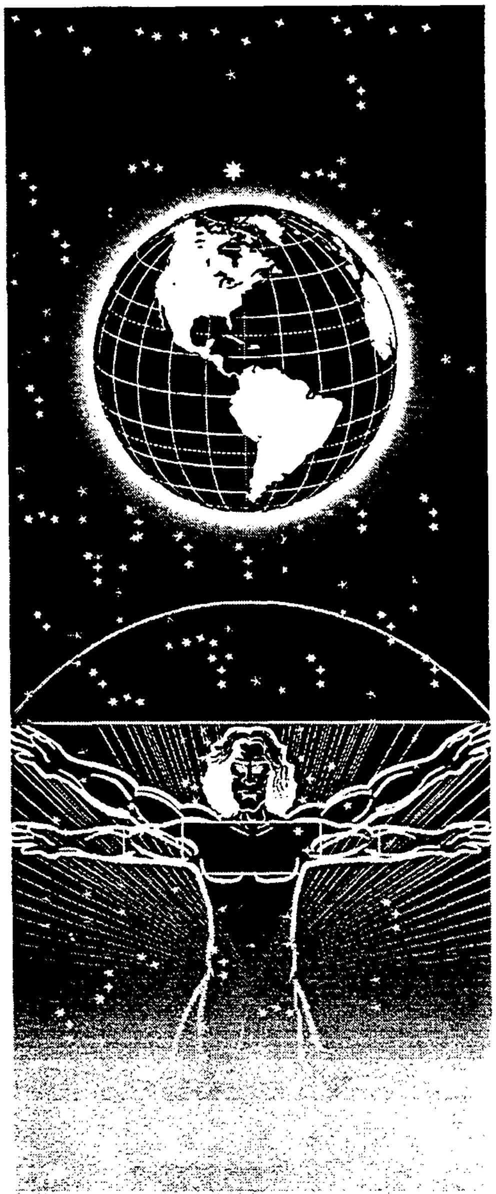
AL/HR-TR-1997-0182



UNITED STATES AIR FORCE RESEARCH LABORATORY

COMBAT MISSION TRAINING RESEARCH AT THE 58TH SPECIAL OPERATIONS WING: A SUMMARY

V. Alan Spiker

ANACAPA SCIENCES, INC.
901 Olive Street
Santa Barbara CA 93101

Robert T. Nullmeyer

WARFIGHTER TRAINING RESEARCH DIVISION
6001 South Power Road, Building 561
Mesa AZ 85206-0904

Steven J. Tourville
Denise R. Silverman

HUGHES TRAINING, INC., TRAINING OPERATIONS
6001 South Power Road, Building 561
Mesa AZ 85206-0904

19980915 123

July 1998

Approved for public release; distribution is unlimited.

AIR FORCE MATERIEL COMMAND
AIR FORCE RESEARCH LABORATORY
HUMAN EFFECTIVENESS DIRECTORATE
WARFIGHTER TRAINING RESEARCH DIVISION
6001 South Power Road, Building 558
Mesa AZ 85206-0904

NOTICES

This report is published in the interest of scientific and technical information exchange and does not constitute approval or disapproval of its ideas or findings.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government-related procurement does not in any way obligate the US Government. The fact that the Government formulated or supplied the drawings, specifications, or other data, does not license the holder or any other person or corporation, or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report has been reviewed and is approved for publication.

ROBERT T. NULLMEYER
Project Scientist

DEE H. ANDREWS
Technical Director

LYNN A. CARROLL, Colonel, USAF
Chief, Warfighter Training Research Division

Please notify AFRL/HEOP, 2509 Kennedy Circle, Building 125, Rm 158, Brooks AFB TX 78235-5118, if your address changes, or if you no longer want to receive our technical reports. You may write or call the STINFO Office at DSN 240-3877, or Commercial (210) 536-3877. E-mail address is shirley.walker @platinum.brooks.af.mil.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
July 1998

3. REPORT TYPE AND DATES COVERED
Final - September 1995 to June 1997

4. TITLE AND SUBTITLE

Combat Mission Training Research at the 58th Special Operations Wing: A Summary

5. FUNDING NUMBERS

C F41624-95-C-5011
PE 62202F
PR 1123
TA B2, B3
WU 06, 01

6. AUTHOR(S)

V. Alan Spiker, Robert T. Nullmeyer, Steven J. Tourville; Denise R. Silverman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Anacapa Sciences, Inc., 901 Olive Street, Santa Barbara CA 93101
HTI-Training Operations, 6001 S. Power Road, Bldg 561, Mesa AZ 85206-0904
Air Force Research Laboratory, Human Effectiveness Directorate,
Warfighter Training Research Division,
6001 S. Power Road, Bldg 558, Mesa AZ 85206-0904

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Training Research Division
6001 South Power Road, Bldg 558
Mesa AZ 85206-0904

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AL/HR-TR-1997-0182

11. SUPPLEMENTARY NOTES

Air Force Research Laboratory Technical Monitor: Dr Robert T. Nullmeyer, (602) 988-6561; DSN 474-6561

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This report summarizes three empirical studies conducted at Kirtland Air Force Base during 1995-1997. The first study examined the relationship between crew resource management (CRM) processes and mission performance for MC-130P Combat Shadow crews who were receiving annual simulator refresher training. Using independent assessments of process and performance, a strong, positive correlation ($r = .86$) was observed between CRM effectiveness at the crew-level and their performance during a simulated tactical mission. A strong association between the quality of a crew's mission planning activities and subsequent mission performance ($r = .60$) was also observed. A second study investigated human factors characteristics of an aerial gunner/scanner simulator (AGSS) recently installed at the 58th Special Operations Wing. The AGSS is a virtual reality (VR) training device that uses a CRT-based, helmet-mounted display and a three degree-of-freedom motion base to train rotary-wing gunners and scanners. A usability assessment by 11 aerial gunner instructors showed that while the devices's VR properties have enormous training potential, the device's human factors aspects need improvement, including the CRTs, head tracker, fitting procedures, and cables. A third study explored the impact of networked simulation on combat mission training. Ninety-nine crewmembers participating in nine networked training exercises were surveyed following training in which MH-53J, MH-60G, TH-53A, and MC-130P weapon system trainers were linked. Survey results strongly support the value of networked training in such areas as multiship tactics, aerial refueling operations, formation flight, situation awareness, and mission team coordination. Areas in need of improvement include establishing training objectives, incorporating emergency procedures into the scenario, and leveling the task demands across crew positions and weapon systems. The report concludes by discussing four recurring themes and four high-payoff areas for follow-on research.

14. SUBJECT TERMS
Aircrews; Annual Refresher Training; ART; CMT; Combat mission; Combat Mission Training; Crew Resource Management; CRM; Special Operations Forces; Team behaviors; Team performance; Teams; Training; Training devices

15. NUMBER OF PAGES
60

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION ABSTRACT
Unclassified	Unclassified	Unclassified	UL

CONTENTS

Page

CONTENTS (Concluded)

FIGURES

TABLES

PREFACE

The research effort documented in this report was performed for the Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Training Research Division (AFRL/HEA), formerly the Aircrew Training Research Division of Armstrong Laboratory, Human Resources Directorate (AL/HRA) by Anacapa Sciences, Inc. and Hughes Training, Inc. (HTI)-Training Operations, in Mesa, AZ. The effort was conducted under Air Force Contract F41624-95-C-5011 with AL/HRA to provide behavioral research support in the areas of combat mission training and mission rehearsal. Anacapa Science, Inc. served as a subcontractor to HTI on this effort providing scientific and technical support. This contract effort was conducted under Work Unit 1123-B2-06, Aircrew Training Research Support; the Laboratory Contract Monitor was Mr Daniel H. Mudd and the Laboratory Technical Contract Monitor was Dr Robert T. Nullmeyer. Inhouse research was conducted under Work Unit 1123-B3-01, Special Operations Forces Aircrew Training and Mission Preparation Research.

We would like to acknowledge several individuals who contributed generously of their time and support to the projects summarized in this report. We would first like to thank LTC Ed Reed who was the Commander of the 58th Training Support Squadron (58 TRSS) throughout much of the duration of these projects. He provided us access to the people and the state-of-the-art technology associated with the 58 TRSS. We would also like to acknowledge Mr John H. Fuller, Jr., HTI's program manager, for his guidance and encouragement throughout these efforts. We would especially like to thank the HTI and Air Force instructors and support staff at the 58th Special Operations Wing and all the participating Special Operations Forces aircrews--their cooperation, input, and tolerance made these projects possible.

COMBAT MISSION TRAINING RESEARCH AT THE 58TH SPECIAL OPERATIONS WING: A SUMMARY

INTRODUCTION

This report summarizes behavioral research accomplished at Kirtland Air Force Base (KAFB) NM using the 58th Training Support Squadron's (58 TRSS), state-of-the-art simulation complex which is capable of supporting combat mission training (CMT) and mission rehearsal (MR). The opportunity for this on-site presence was made possible through a research partnership between the 58 TRSS's parent organization, the 58th Special Operations Wing (58 SOW) and the Air Force Research Laboratory's Warfighter Training Research Division (AFRL/HEA).

The 58 SOW's mission involves the training and combat qualification of aircrews in the areas of Special Operations and Combat Search and Rescue (CSAR). Assigned to the Air Education and Training Command (AETC) in 1993, the Wing provides mission qualification (MQ), annual refresher training (ART), and upgrade qualifications (e.g., becoming an instructor pilot or a flight examiner) for a variety of Air Force Special Operations Command (AFSOC) weapon systems. These include the MH-53J PAVE LOW III, M/HH-60G PAVE HAWK, MC-130P COMBAT SHADOW, and MC-130H COMBAT TALON II (Reed & Selix, 1995).

The overarching goals of the researcher team were to conduct scientifically sound and operationally relevant field research, and to develop written technical products of immediate value to the United States Air Force (USAF) and the 58 TRSS.

In the next three sections, we summarize the three major empirical studies that were conducted: (a) measuring crew resource management (CRM) behaviors during CMT; (b) conducting a human factors evaluation of the Aerial Gunner Scanner Simulator (AGSS); and (c) evaluating perceived benefits of integrated networked training using the Special Operations Forces Network (SOFNET). Each section summarizes operational problems and associated technical issues that gave rise to the research, objectives and hypotheses that were tested, methods used to conduct the study, major results that were obtained, implications for training, and technical products issued by the research team on this topic. Note that these sections only give synopses of the research that was conducted; in each case, the reader is referred to full-length reports that provide complete descriptions of the technical work.

The final section of this report discusses a number of high-payoff, follow-on research and development (R&D) research data, observations, and feedback from 58 TRSS personnel. We begin by discussing four recurrent themes from our research at KAFB. These are the need to: (1) review training requirements and performance objectives, (2) better integrate ground training and flightline operations, (3) improve student performance data, and (4) demonstrate effective training strategies. We close with a discussion of future research possibilities: revising and assessing MC-130P CRM course curricula, extending our CRM methodologies to rotary-wing aircraft, developing measures and methods for better serving the information needs of training managers and decision makers, and identifying strategies for training mission planning.

CRM AND COMBAT MISSION TRAINING

This section summarizes the study of the CRM process that we performed at the 58 TRSS from January to October 1996 in the context of normal ART for SOF MC-130P aircrews. The overall goal of the study was to determine if crews who engaged in more effective CRM processes would also exhibit superior mission performance. The basic premise, along with a number of offshoots, was confirmed in the empirical data.

Background

Operational Context

In the 1970s, systematic reviews of airline accident data revealed that crewmembers who were clearly skilled in their individual tasks committed numerous errors. NASA surveys and research suggested that aircrew performance, as measured by accident incidence, could not simply be predicted from individual crewmember proficiency (Lauber, 1987). Analysts surmised that the performance of these team coordination tasks made the successful cockpit crews more effective. Such behaviors were collectively referred to as CRM, and both the airline industry and the military launched a series of CRM training initiatives in the 1980s to improve flight safety and reduce pilot errors (Spiker, Tourville, Silverman, & Nullmeyer, 1996). Focusing initially on the individual pilot, CRM has evolved to include other crewmembers as well as the entire combat mission team (Andrews, Bell, & Nullmeyer, 1995).

Air Force Instruction (AFI) 36-2243, Cockpit/Crew Resource Management Training (AFI 36-2243, 1994) establishes the requirement that each major command (MAJCOM) measure and train CRM to: (a) maximize operational effectiveness and combat capability, (b) prevent accidents and incidents, and (c) improve all forms of training efficiency. With regard to CRM content, the AFI dictates that the core curriculum covers eight concepts: situational awareness (SA); group dynamics; effective communications; risk management and decision making; workload management; stress awareness and management; mission planning, review, and critique strategies; and human performance.

As a MAJCOM, AFSOC must comply with the Air Force requirement (AFI 36-2243, 1994) during each phase of training, including ART. The document broadens the measurement domain of CRM-related performance: "Although CRM programs are mandatory for aircrew members and have historically been geared toward the operational flying environment, the potential exists to adapt fundamental program principles to *any task or functional area requiring cooperative or interactive time critical efforts* (p. 1)." This latter stipulation opens the door to examine team coordination behaviors from multiple methodological perspectives.

In addition, the AFI stipulates that two methods for CRM program evaluation are acceptable: (a) over-the-shoulder observation by evaluators, qualified line observers, and MOST [mission-oriented simulation training] instructors to collect skill performance data, and (b) the conduct of "derivative" (before-after) crew surveys. This stipulation is quite significant from a research standpoint as it expands the range of acceptable measurement options over and above the three traditional metrics: inflight assessments, mishap tracking, and student critiques. This is particularly important for measuring performance in a simulator for which accident data and check-ride ratings are not applicable (Spiker et al., 1996).

In ART for SOF MC-130P aircrews, CMT is a combination of CRM academics and simulator training, technical training (e.g., emergency procedures (EPs) and systems), and combat tactics (D. Wilson, personal communication, September 28, 1995). The latter covers a range of events, such as threat recognition, expendables deployment, mission planning, air refueling (AR), low-level navigation, and covert insertion and extraction. It is this added emphasis on developing or exercising tactical skills that makes the study of military teams so different from its commercial counterpart, necessitating the development of customized measurement procedures and data collection instruments.

At the squadron level, there are two problems in implementing an effective CRM training program within existing USAF guidelines. First, governing Air Force directives do not specify training objectives for CRM as an ART training event. Consequently, the CRM scenarios that guide CMT in advanced simulators do not have any accompanying CRM training objectives. Second, in the MQ precursor to ART, CRM is taught in academics to all crewmembers using the generic AFSOC CRM workbook. Taking a "one-size-fits-all" approach to CRM training, AFSOC guidance does not address the use of simulation for CRM training because some weapon systems do not have simulators for either MQ or ART. This leaves it up to the Formal School to decide whether CRM training is a good use of scarce and expensive simulator resources. The AFSOC approach also does not address aircraft- or mission-specific requirements.

Despite several attempts to demonstrate the effectiveness of generic CRM training, there are no data to our knowledge to support the value of this approach. Informal observations of aircrews receiving academic and simulator training indicate that, while some general topic matter is considered useful, there is widespread dissatisfaction with the lack of tactically relevant, team coordination behaviors that would be useful within a specific weapon system or mission type. It was in this context that our CRM research was designed.

Research Issues

There is a presumed importance of CRM for achieving mission success. However, there have been few empirical studies that directly link effective CRM with mission outcome. One such study was Povenmire, Rockway, Bunecke, and Patton's (1989) observation of B-52 aircrews as they flew a complex, tactically realistic, mission scenario in a high-fidelity weapon system trainer (WST). The scenario entailed conducting a long-range bombing mission requiring the penetration of enemy threats, accurate dropping of bombs, and intricate navigation and maneuvers. Highly trained CRM evaluators assessed crew coordination and mission performance, with separate sets of raters used for each measure. Three measures of mission performance were taken: bombing accuracy, threat avoidance, and technical skill.

The primary analysis assessed the correlation between overall aircrew coordination ranking and the crew's mission performance ranking where a strong, positive relationship ($r = .84$) was obtained. Povenmire et al. (1989) then compared the experts' ratings of mission performance with the individual mission outcome factors, revealing that the experts primarily used bombing accuracy to make their overall judgments of mission performance. Similarly, overall crew coordination was compared with individual coordination dimensions. This analysis showed that raters based their judgments of overall crew coordination on four dimensions—practicing inquiry and advocacy, avoiding distractions, distributing workload, and resolving conflicts.

A primary objective of our research was to see if we could replicate Povenmire et al.'s (1989) observation of a strong, positive, CRM-mission performance association within the context of a SOF MC-130P simulated tactical mission during ART. Two activities were performed in advance of the present study. The first was to expand the scope of CRM, from the "soft" focus used by the airlines to a more technical, tactics-based orientation. In this regard, it was clear during our front-end analyses that the traditional, organizational-centered conception of CRM would have limited applicability to the complex, turbulent CMT environment. This view was substantiated during pilot tests, observations, and interviews conducted with SOF subject-matter experts (SMEs) in the early stages of this project (Spiker et al., 1996). Based on these analyses, we adopted the "tactical team resource management" concept, or T^2RM, as a logical extension of the CRM approach discussed above. As the name suggests, T^2RM embraces three elements that distinguish it from CRM: (a) an enhanced focus upon tactical skills, (b) a clear delineation of a combat mission team, and (c) an emphasis on managing a multitude of diverse resources within a dynamic environment.

A second activity that we performed was the development of a detailed measurement model for T^2RM (see Figure 1). The model lays the conceptual foundation for using multiple "data hooks" to capture crewmember behaviors and interactions with other team players, monitoring tactically relevant behaviors, and measuring behaviors across the entire range of mission events. A complete description is in Spiker et al. (1996) and Spiker, Silverman, Tourville, and Nullmeyer (in press).

The boxes and arrows in Figure 1 flow from left to right, reflecting an implicit timeline of SOF MC-130P ART activities. The white boxes in the interior represent ART elements that must be measured during the course of training to establish links between T^2RM and crew mission performance. The boxes on the periphery indicate the instruments that were used to collect attitude, performance, process, and outcome data from crewmembers and aircrews.

Our primary analyses focused on the T^2RM and Team Mission Performance/Mission Outcome modules, where following Povenmire et al. (1989), data were collected by two independent observers using different instruments (i.e., the Team-Mission Observation Tool, or T-MOT, and the Team-Mission Performance Tool, or T-MPT). The overall T^2RM process is subdivided into five main subprocesses: situation awareness (SA), function allocation (FA), time management (TM), tactics employment (TE), and command-control-communications (C3). These five were selected for study based on their: judged relevance to the AFSOC mission environment, appropriateness to the high levels of experience and motivation of many MC-130P aircrews, applicability to CMT, and amenability to measurement by outside observers. Where possible, we attempted to identify subprocesses that made contact with the CRM dimensions previously studied by other researchers (Spiker et al., 1996).

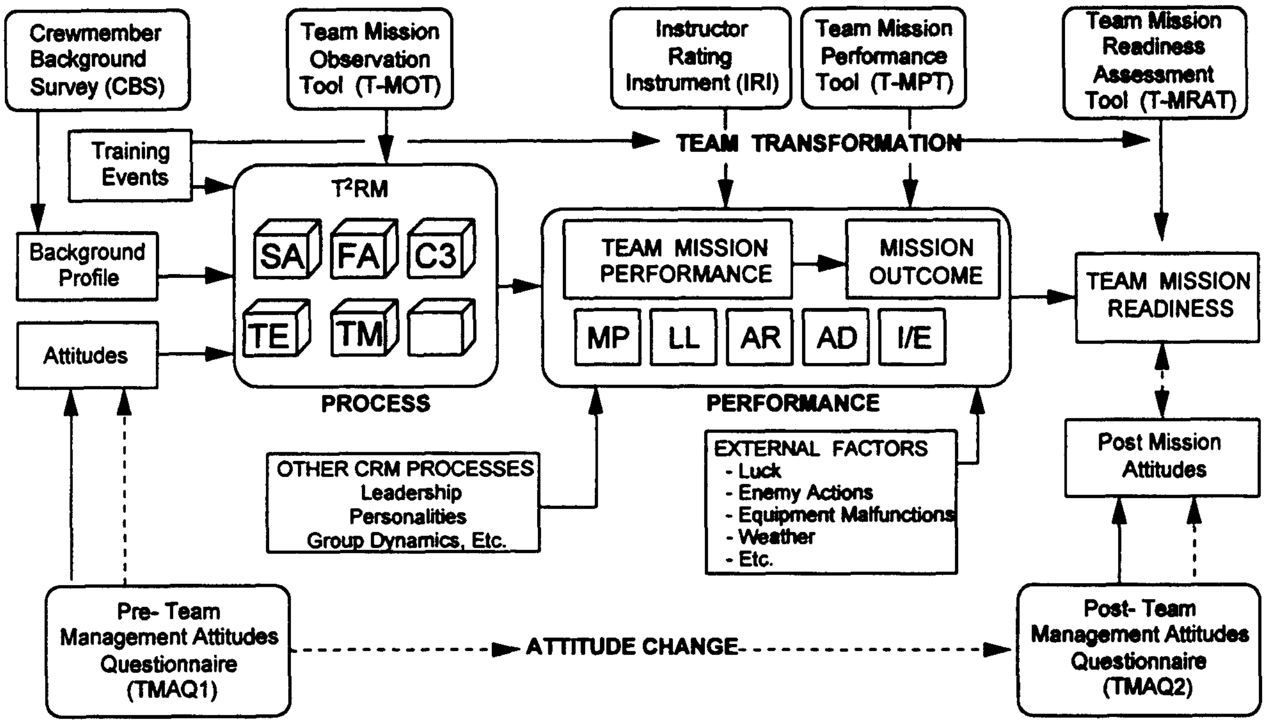


Figure 1. T²RM Measurement Model (Silverman, Spiker, Tourville, & Nullmeyer, in press).

In addition, the Other CRM Processes module is depicted in the gray-shaded box beneath the T²RM module. These processes include such traditional process variables as group cohesiveness, personality, group dynamics, and leadership. While these factors may have some influence on team mission performance, they were not measured in our research.

The output of the T²RM module feeds into the Performance module. By team mission performance, we mean those indices that directly result from successful (or failed) execution of important T²RM processes. In our research, team mission performance is reflected in such indices as quality of mission briefings, completeness of navigation chart(s), and instructor-supplied ratings of how well the team as a whole and individual crewmembers executed each mission phase. (In the MC-130P, this team includes the Aircraft Commander (AC), the Copilot (CP), the Left and Right Navigators (LN, RN), the Flight Engineer (FE), and the Communication Systems Operator (CS0).) It is these indices which we wished to correlate with T²RM process behaviors as they are less susceptible to external factors than are "bottom line" mission outcome variables (e.g., success versus failure). Also, instructors tend to modify the simulated environment during training scenarios to ensure mission completion, which tends to reduce performance-driven variability of outcome across crews.

Questions/Objectives

This study addresses 10 questions:

1. *Are team coordination processes strongly and positively related to mission performance?* Crews exhibiting superior coordination behaviors should perform better during the mission than those who do not. This relationship is the fundamental tenet of our research, as a significant finding gives us "permission" to probe the data further, to test the remaining hypotheses.

2. *Are the five T^2RM subprocesses—FA, TM, SA, TE, C3—differentially related to performance across mission phases?* Since T^2RM is not likely to be a single entity, we should not expect every T^2RM subprocess to be significantly correlated with every phase of mission performance. Rather, only one or two T^2RM subprocesses may predict performance in a given phase.

3. *Is the quality of mission planning positively related to mission performance?* Crews who do better mission preparation should be more successful in the mission than those who do not. While a positive relationship between planning and performance is at the heart of all military planning, there is surprisingly little empirical data to substantiate this assumption.

4. *Are there measurable aspects of team structure that are related to crew effectiveness?* The more "effective" crews should have structural components in common that are not present in effective crews. These might include combined crewmember experience in the area of operation (AO), hours flown together as a crew, or the manner in which the crew is organized as a team.

5. *Is there a positive relationship between crew perceptions of T^2RM effectiveness and their perceptions of mission performance?* Crews should be able to report how well they executed T^2RM upon completing the mission. Crews who perceive their coordination to have been successful should also report having performed better during the mission.

6. *Is the T^2RM effectiveness of all crew positions equally important in determining overall team mission performance?* Based on reports by SMEs, we expect the role of the LN to be larger than those of other crew positions.

7. *Does overall T^2RM for any particular crew position have a stronger relationship to overall crew T^2RM or are they approximately the same?* Just as some crew positions might differentially impact mission performance, their coordination behaviors might also differentially impact overall T^2RM. We expect the AC, LN, and RN to have a bigger impact than the FE, CP, or CSO.

8. *Will the impact of crew position T^2RM effectiveness on team performance be similar across each mission phase?* Given the diversity of mission events performed by the MC-130P, different crew positions may "step forward" to differentially impact mission phase T^2RM.

9. *Is the relative impact of crew position on overall T^2RM the same across each mission phase?* While it is possible that all crew positions are equally vital in promoting the "emergence" of an effective team, MC-130P operations may dictate that certain crew positions play a larger role than others in accomplishing various mission events.

10. *Are there differential impacts of specific T^2RM subprocesses by crew position?* For example, a good LN may rate high on SA and TM, whereas TE, FA, and C3 may not be as important. On the other hand, an effective CSO may rate high on TM and C3, but lower on TE, FA, and SA. While such determinations can become rather involved (six crew positions by five T^2RM subprocesses), they will help us specify the content of future training interventions.

Method

Participants

Eleven MC-130P SOF aircrews (67 crewmembers total) were observed during Day 5 of ART. Though small, the sample represents 26% of the total population of SOF MC-130P crews who go through ART each year. Crews came from several different operational squadrons and had an average of 3,056 flight hours. Six crewmembers undergo ART in this aircraft: AC, CP, FE, LN, RN, and CSO.

Equipment

The MC-130P WST is a six-degree-of-freedom, high-fidelity simulator. This device uses a CompuScene V image generation (IG) system, a fully correlated infrared detection system, a digital radar landmass system, state-of-the-art navigation systems, and out-the-window displays. An electronic warfare (EW) database contained preprogrammed ground threats that emulated actual emitters including visual and audio depictions on the crew's threat warning receiver system.

Mission Scenario

The training mission, composed of five phases, took approximately seven hours to complete. The first three hours were devoted to **mission preparation** (MP) which took place in a small briefing room outside the simulator bay. The mission was to support recovery of injured

Army personnel by providing AR tanker coverage for two MH-53J Pave Low helicopters that would be transporting the injured to a field hospital. The crew was to airdrop a special forces team at a drop zone (DZ) who would prepare the evacuees for transport at the landing site for the transload operation. A secondary tasking was to transport a flag officer and his staff by covertly airlanding at the field hospital.

Following the initial briefing, crews were given weather reports and charts of the tactical areas, operations and communications details, Rules of Engagement (ROE), and Order of Battle (OB) threat data. The crews were to integrate these materials into an executable mission plan. The period concluded with briefings by the crew on the details of that plan. After lunch, the crews entered the WST with the products generated during MP. The crews then executed the other four mission phases: **low-level** (LL) navigation under night vision goggle (NVG) and terrain masking conditions, an **air refueling** (AR) constrained by altitude and threats, an **airdrop** (AD) to a "blind" DZ, and a covert **infiltration/exfiltration** (I/E).

Design

There were five central features to the study design. We collected data using a *naturalistic observation* of a CMT mission scenario already in place rather than an experimental manipulation. *Independent* collection of coordination data from one researcher and performance data from a second researcher avoided inflated correlations when both sets of measures come from the same rater. The study's *robust measures* of team mission performance and coordination focused on behaviors that were collectible, variable across crews, and operationally relevant. We used a *multi-measure, multi-method* mix of subjective and objective variables to assess T^2RM and mission performance. *Behaviorally anchored rating scales* (BARS) served as criterion standards to structure ratings for both T^2RM and mission performance.

Procedure

During MP, the SME-researcher recorded notable crew behaviors and helped present the training materials, role play outside agencies, and provide mission debrief support. During mission execution, he observed crews from an intercom station outside the WST and situated in front of four instructor/operator station (IOS) screens which repeated the instructor inputs from inside the WST.

The SME-researcher recorded T^2RM process data on the *Team-Mission Observation Tool* (T-MOT). The T-MOT is a 20-page booklet that structures SME observations (Spiker, et al., 1996) by using a "record-by-exception" measurement approach to capture instances of extreme crew coordination behaviors. The T-MOT was divided into five subsections, one for each subprocess. Each subsection contained a customized set of questions, YES/NO checklist items, space for recording notable behaviors, and a 5-point rating (1 = poor, 5 = good) of the crew on that subprocess and that phase. Figure 2 shows an item from the TM section of the MP phase.

Time Management (TM): Involves the ability of the combat mission team to employ and manage limited time resources, so that all tasks receive sufficient time to be performed correctly, and critical tasks are not omitted.

1.0	An end-mission planning time should be indicated up front - most likely by an emergent "leader."		
	a.	Did any crewmember indicate the need for an end-mission planning time? .... (Explain) ________	YES / NO
	b.	Was that time noted by all other crewmembers? .... (Explain) ________	YES / NO
	c.	Did any crewmember designate activities to establish a proper balance between their own authority, time available, and crewmember participation? .... (Explain) ________	YES / NO
	d.	Was adequate mission preparation time allocated for a comprehensive pre-mission briefing? .... (Explain) ________	YES / NO

Figure 2. Example Item from the T-MOT (Spiker et al., 1996).

A second researcher collected performance data using the *Team-Mission Performance Tool (T-MPT).* The T-MPT is a structured method for rating the quality of products created during MP and provides ratings of a crew's mission performance during the LL, AR, AD, and infiltratration/exfiltration (I/E) phases. Figure 3 depicts a scale that was used to rate mission flight charts developed by the navigators and the pilots during MP.

FLIGHT CHARTS

1	2	3	4	5
- Poor. - Incomplete data. - General lack of documentation. - General quality of preparation is poor.	- Marginal. - Insufficient or inaccurate documentation. - Unaccounted for discrepancies between LN, RN, and CP charts. - Deviation plan minimally prepared. - Marginal quality.	- Adequate. - Threats plotted. - Most threat rings plotted. - Deviation plan clearly drawn. - Appropriate altitude considerations made. - Required checklist annotations made.	- Outstanding. - Threat rings plotted. - Deviation plan clearly drawn and visible for NVG conditions. - Appropriate altitude and terrain considerations made and explicitly represented in the deviation plan.	- Exceptional. - Threat contour shading provided. - Deviation plan and threat information highlighted for NVG conditions. - Documentation in excess of minimum requirements. - Threat labels.

Figure 3. Flight Chart Rating Scale from the T-MPT (Spiker et al., 1996).

During mission execution, the researcher monitored crew communications, flight path, and threats from a separate room designed for training the CSO. This room also permitted printing of key mission performance data, such as the aircraft's ground track at each waypoint, which was used to aid in assigning the behaviorally anchored performance ratings (e.g., planned versus actual ground track). At the study conclusion, the two researchers independently reviewed their data packets and rank ordered the 11 crews in terms of overall T^2RM and mission performance, respectively.

Major Results

1. Are team coordination processes strongly and positively related to mission performance? The answer to the primary question of our study is a definite "Yes." The correlation between overall T^2RM process and overall mission performance is .86 ($t = 6.143$, $df = 9$, $p < .05$, two-tailed), which is comparable to that obtained by Povenmire et al. (1989). The scatterplot in Figure 4 depicts the linear relationship between these two variables. As is evident from the figure, the poorest performing crews did indeed have the lowest overall T^2RM process ratings whereas the best performing crews had the highest process ratings.

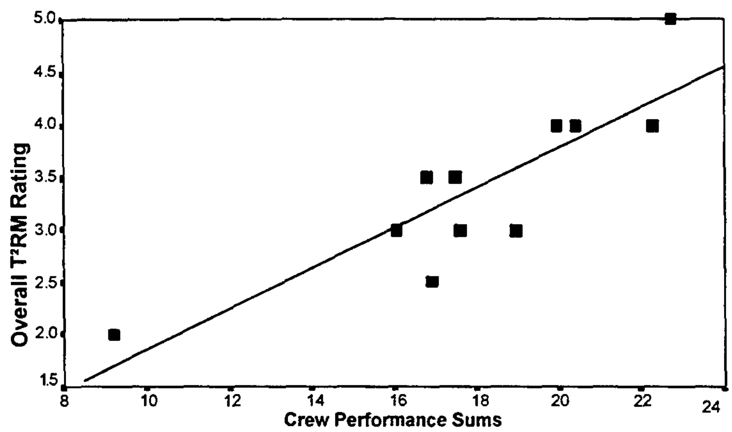


Figure 4. Scatterplot Between Overall Crew Performance and Overall Crew T^2RM (Spiker et al., in press).

2. *Are the five T^2RM subprocesses differentially related to performance across mission phases?* The answer to this question is also "Yes." First, analysis of correlations between T^2RM subprocesses and total mission performance showed that four of the T^2RM subprocesses—SA, TE, TM, and FA— were significantly related to overall performance, with correlations ranging from .75 to .83. Crew level C3 ($r = .08$) was not. A more detailed probing revealed that different T^2RM subprocesses played larger roles in different mission phases indicating that T^2RM is a multidimensional entity. During MP, the SA and TM subprocesses were most important (as evidenced by higher correlations). During the four phases of mission execution, the following subprocesses were dominant: LL--FA, AD--TM, AR--TE, and I/E--SA and TM. Once again, C3 was not significantly associated with performance in any of the phases.

3. *Is the quality of mission planning related to mission performance?* The answer is most definitely "Yes," as the SME's ratings of crew coordination process during mission preparation were significantly correlated with the second researcher's ratings of overall crew mission performance ($r = .78$; $t = 3.74$, $df = 9$, $p < .005$). The overall rating of MP *products* was also a significant predictor of crew performance during the four mission execution phases, with $r = .60$ ($t = 2.25$, $df = 9$, $p<.05$). As a third index of mission planning quality, each crew was scored on the extent to which it exhibited positive or negative behaviors on 12 dimensions of MP effectiveness. Dimensions included using personnel effectively, establishing a timeline, holding

interim briefings, checking information sources for accuracy, and so on (Spiker, Tourville, Silverman, & Nullmeyer, in press). Crews exhibiting positive behaviors on more of the MP effectiveness dimensions performed significantly better during the mission, with $r = .71$ ($t = 3.02$, $df = 9$, $p < .01$).

4. *Are demographic variables and team structure related to crew effectiveness?* The answer to this question is a marginal "Yes." Due to the small number of crews in this study, it was difficult to identify particular demographic variables that were statistically related to effectiveness. Nevertheless, two variables exhibited discernible trends. First, squadron affiliation was relevant, as crews from one squadron achieved the lowest effectiveness scores whereas the highest marks came from another squadron. Second, the most effective crews had a distinct "hub-and-spoke" structure that promoted interactions and information sharing whereas the weaker crews "never came together" over the course of Day 5 ART (Spiker, et al., in press).

5. *Is there a positive relationship between crew perceptions of T^2RM effectiveness and their perceptions of mission performance?* Following the mission, crews completed a questionnaire that asked for self-ratings of crew coordination and mission performance. A strong positive correlation was again found, $r = .86$ ($t = 6.025$, $df = 9$, $p < .05$, two-tailed), indicating that members who felt their crew did well on the mission also believed their crew coordination was good. Although these two data sources are not independent, they nonetheless provide secondary validation of the link between T^2RM and combat mission performance.

6. *Is the T^2RM effectiveness of all crew positions equally important in determining overall team mission performance?* Given the varied duties required of MC-130P crewmembers, it is not surprising that the answer to this question is "No." For each crew position, correlation between overall T^2RM and overall mission performance was computed. They are displayed in Figure 5. The T^2RM effectiveness of the LN, CP, and AC positions was more strongly related to mission performance than were the other three crew positions.

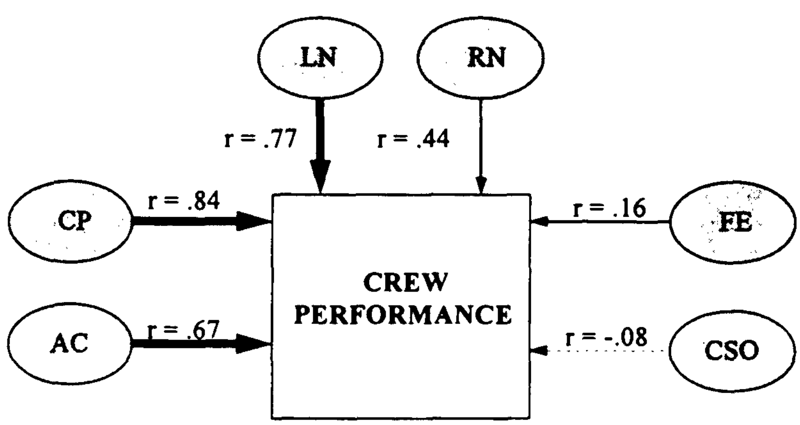


Figure 5. Correlations between Crew Position T^2RM and Mission Performance (Silverman et al., in press).

7. *Does overall T^2RM for any particular crew position have a stronger relationship to overall crew T^2RM?* This question parallels the previous one, where we consider the relationship of crew position T^2RM to overall T^2RM rather than mission performance. Again, the answer to this question is a resounding "Yes." Four crew positions--AC, CP, LN, RN--were highly predictive of overall T^2RM, with correlations ranging from .64 (RN) to .89 (LN). FE and CSO correlations ($r = .41$ and .20 respectively) were significantly weaker than the group average.

8. *Will the impact of crew position T^2RM effectiveness on team performance be similar across each mission phase?* This question is an extension of Question #6, in which we test whether particular crew positions are more or less important to performance in a particular mission phase. The answer is again "Yes," where the three crew positions highly predictive of mission performance across each phase are the AC, CP, and LN. Correlations for these crew members were uniformly highly positive, ranging from .72 (LN during MP) to .39 (LN during AD). When considering crew position impact across phases, detailed analyses revealed that the FE had less influence than the group during MP, AD, AR, and I/E. The CSO had less impact during MP and I/E whereas the RN had less impact during the LL and AD phases.

9. *Is the relative impact of crew position on overall T^2RM the same across each mission phase?* This question is a logical extension of Question #7, in which we examined the effects of each crew position's T^2RM on crew-level T^2RM by mission phase. The answer here is clearly "No," as T^2RM importance varies with crew position and different profiles are obtained across the mission phases. The impact of the CP's T^2RM was significantly greater than the group average in four phases: MP, LL, AR, and I/E. Similarly, the LN's T^2RM impact was greater than the group's in AD, AR, and I/E. The AC's impact was also greater than the group's in the I/E phase. Finally, only during MP does the impact of the FE's T^2RM match that of the group.

10. *Are there differential impacts of specific T^2RM subprocesses by crew position?* This question asks whether there is a different functional relationship between the T^2RM subprocesses and crew T^2RM for each crew position and do these relationships vary across the mission phases. Given the previous findings, it is not surprising that the answer is "Yes." A complete answer requires an analysis of six correlation tables, one for each crew position. That analysis, along with a summary of the exceptional T^2RM behaviors by crew position and mission phase, is presented in Silverman et al. (in press).

The correctional data clearly show that the importance of a T^2RM subprocess varies across crew position. For example, SA was the dominant subprocess for the AC whereas C3 was for the CSO. SA subprocess observations revealed that one of the AC's notable behaviors is that he cross-checks everyone's work. In addition, the relative importance of the five T^2RM subprocesses for a given crew position was not stable across mission phase. For example, TM was more important for the RN during MP whereas TE was dominant in LL. Importantly, this combination of quantitative and qualitative analyses will aid in an initial identification of content areas to include in a revised CRM course curriculum.

Implications for Training

The first area involves establishing training objectives and associated behavioral standards for CRM. These objectives should be platform-, mission-, and task-specific, and should include specifications for: (a) which crewmembers are to perform which tasks, (b) the criterion level for task performance, and (c) the scope of the objective (e.g., Does it only apply to ART, or to MQ as well?). If criterion-based objective standards for performance existed, we could more reasonably state in simple terms the underlying reasons for a crew's poor performance.

The second area is the need for improved CRM training methods to better satisfy training objectives where it is known that ART crews perceive a lack of relevance in their CRM training. Following an initial academic training period, students do not perform any CRM-related skills and behaviors until the CMT scenario on the last day of ART. Our observations support changing the overall CRM training method in ART to include: specific team coordination training objectives, a revamped CRM curriculum, a modular approach to simulator utilization, and a flexible approach to CMT scenario construction (Spiker et al., in press).

The third area involves improving feedback to student crews concerning CRM performance and to the training organization concerning effectiveness. Despite wide variation in crew CRM and mission performance, the content of instructor debriefs was fairly constant. Greater top-level USAF management support for the importance for CRM debriefing is needed as are improved tools (e.g., debrief checklist) to support instructor measurement of tactical proficiency, reinforcement of good CRM behaviors, and critique of poor CRM behaviors.

The fourth area is MP training. Our data show that comprehensive MP is an extremely important factor to ensure mission operations success, but this activity is downplayed at the ART level where it is viewed as separate from the actual simulator training period. To that end, we recommend better incorporation of advanced planning materials and semi-automated systems into the MP process, development of methods for students to structure their own "team" coordination activities that promote effective planning, and academic reviews of important planning principles.

Technical Products

The following ten technical products were delivered pursuant to work on CRM:

Nullmeyer, R.T., Silverman, D.R., Tourville, S.J., & Spiker, V.A. (in press). Data collection instruments used to measure tactical team resource management (T^2RM) behaviors. (AFRL/HEA-TP-1997-xxxx). Mesa, AZ: Air Force Research Laboratory, Human Effectiveness Division, Warfighter Training Research Division.

Silverman, D.R., Spiker, V.A., Tourville, S.J., & Nullmeyer, R.T. (1996, November). A combat team performance model: Development and initial application. *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Silverman, D.R., Spiker, V.A., Tourville, S.J., & Nullmeyer, R.T. (in press). *Crew position processes effect on team performance during combat mission training.* (AL/HR-TR-1997-0136). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Silverman, D.R., Spiker, V.A., Tourville, S.J., & Nullmeyer, R.T. (1997, November). Team coordination and performance during combat mission training. *Proceedings of the 19th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Spiker, V.A., Nullmeyer, R.T., Tourville, S.J., & Silverman, D.R. (1997, September). Effects of mission preparation on crew combat mission training performance. Paper presented at the *41st Annual Meeting of the Human Factors and Ergonomics Society*, Albuquerque, NM.

Spiker, V.A., Tourville, S.J., Silverman, D.R., & Nullmeyer, R.T. (in press). *Tactical team resource management effects on combat mission training performance*. (AL/HR-TR-1997-0137). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V.A., Tourville, S.J., Silverman, D.R., & Nullmeyer, R.T. (1996, November). *Team performance during combat mission training: A conceptual model and measurement framework.* (AL/HR-TR-1996-0092). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Tourville, S.J. (1997, April). Identification of key behaviors for effective team coordination in Special Operations Forces combat mission training. Paper presented at the *Ninth Symposium on Aviation Psychology*, Columbus, OH.

Tourville, S.J., Spiker, V.A., Silverman, D.R., & Nullmeyer, R.T. (1996, November). An assessment methodology for team combat mission training and rehearsal. *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

AGSS HUMAN FACTORS EVALUATION AND INITIAL IMPACT ASSESSMENT

During 1996, the HTI research team conducted a human factors investigation and an initial MQ training impacts assessment of the AGSS. The results of the AGSS human factors evaluation indicated that there were several aerial gunner/scanner (AG/S) skill areas that could be effectively trained using the device, and though some usability problems were reported, none were insurmountable. In our initial training impact assessment, we observed some very innovative AGSS training given the current configuration of the device (as a standalone trainer) and delays in curriculum development. Students have also been quite receptive to AGSS training. However, we had difficulties in determining any direct impacts of the AGSS on flightline training, aside from reductions in AG flying hours dictated by the syllabus of instruction (SOI) and concurrent with other curriculum changes. The study design features and findings of the twofold effort regarding the AGSS are described below, with an emphasis on the human factors portion and interjections regarding the training impact assessment where applicable.

Background

Operational Context

In 1990, the 58 SOW received the MH-53J WST and its associated database generation system (DBGS) and mission planning system. As the first of a series of advanced WSTs, the MH-53J WST underwent an initial operational effectiveness evaluation the following year in the context of a joint Air Force-Army training exercise (Nullmeyer, Bruce, Conquest, & Reed, 1992). Though overall assessments by the MH-53J WST pilots and FEs were quite positive, room for improvement was cited by the AG/Ss who could not rehearse in the 3-seat (2 pilots, FE) device. They described the benefits of full-crew rehearsal and heightened SA, as well as the need for scanners to have their own visual capability and communication/coordination with the cockpit crew during rehearsal.

Based on this feedback, it was recommended that a simulation capability be developed that allows the AG/Ss to view the visual terrain database just as the pilots and FEs do, only from their vantage point at the rear and sides of the aircraft (Nullmeyer et al., 1992). Not only would such a capability permit the training of full crew tasks, it would aid development of key gunnery and scanning skills while reducing demands on the flightline to schedule sorties solely to accomplish AG/S-unique mission training events.

While creating the requirements for an AG/S training simulator, the 58 TRSS explored promising new technologies for incorporation into their formal specification. These included low-cost, small motion platforms; ruggedized, full-color helmet-mounted displays (HMDs); more accurate head-tracking systems; and smaller, high-fidelity IGs.

These operational and training requirements, combined with emerging technologies, gave rise to the AGSS, which was declared ready for training (RFT) in March 1996. The AGSS was designed to provide a low-cost, high-fidelity virtual environment for training MH-53J Pave Low and M/HH-60G Pave Hawk FEs and AG/Ss. Its dynamic virtual environment is presented through a head-tracked, full-color HMD system that can support both daytime and NVG simulation. Using a 3-degree-of-freedom motion system, aural cueing, a powerful 7-channel IG, and digital computation system, the AGSS is capable of simulating air, sea, and terrain scenes; ground and airborne targets; and weapons effects, including tracer path and bullet impact (Reed, 1996).

Because the AGSS represents the first use of virtual reality (VR) for aircrew training, the 58 TRSS requested a human factors assessment before incorporating the system into its impressive cadre of advanced WSTs (Silverman, Spiker, & Nullmeyer, 1996). Four system design aspects were particularly thought to require a formal evaluation:

(1) the modular system architecture which permits reconfigurability between MH-53J and MH-60G training,

(2) an IOS located external to the AGSS main compartment,

(3) operation of the AGSS either in independent (standalone) or integrated mode (networked with a cockpit simulator); and

(4) creation of an immersive training environment afforded by the system's revolutionary IG capability, HMD display of digitally scanned representations of aircraft interior, position-referenced aural cues, and vibratory feedback from the guns.

Research Issues

Within DOD, VR is viewed as an "interactive, usually computer-mediated, experience characterized by a suspension of (or inattention to) disbelief in the unreality of the experience" (DOD, 1994, p. 18). Given the user's large role in the VR experience, it stands to reason that the human factors aspects of the implementation will influence how effective the simulation will be. In this regard, it should be noted VR technology is not a single technology or medium, such as an HMD or a DataGlove™, but is an integrated human-computer system (Thurman & Mattoon, 1991). Consequently, when VR is incorporated into a training device, all aspects of the human-computer system--including input, interface, processing, and output--must be evaluated, both separately and in combination. When assessing the effectiveness of a VR-based training device, it is important to address how: *accurately* the simulation represents the objects of interest, *realistically* the modeled objects behave over time, and *efficiently* the modeled environment is delivered to the user.

Besides the "virtual" aspects of the VR environment, two other essential features are immersion and presence. Immersion is achieved when one or more of the user's senses (typically, vision and audition) are isolated from the surrounding environment and fed only information from a computer simulation such as an HMD (Pimentel & Teixeira, 1993). This feature may have potential drawbacks in training, as it might be difficult for an outside observer (the trainer) to attract the attention of an "immersed" trainee. Presence is the subjective experience of being in one place when one is physically in another (Bailey & Witmer, 1994). This may have a positive impact on VR-based training, as there is evidence to suggest that users can still experience presence while modifying the viewpoint in the virtual environment, thus improving their spatial knowledge of objects in the virtual world (Mowafy & Miller, 1993).

While the potential for VR technology to positively impact CMT is quite real, some human factors-related problems have been highlighted in past implementations that should be assessed. These include eyestrain, blurred vision, and headaches caused by delays in the head-tracking mechanism (DOD, 1994); perceptual problems due to poor resolution in the HMD; disorientation while navigating in the virtual environment; and simulator sickness (Bailey & Witmer, 1994).

Regardless of the method of display or the medium of control, both the user and the machine will process substantial information. Since the processing of both entities is hidden from direct observation, the challenges facing the human factors analyst are greatest here (Kantowitz & Sorkin, 1983). On the machine side, the system must respond sufficiently soon after input so users do not experience any delay that might degrade their performance. On the user side, the goal is to design the system so that user's cognitive processing capacities (attention, short-term memory, long-term memory, etc.) are not overwhelmed.

A human factors analysis must address any anthropometric (i.e., fit and feel) problems posed by wearing an HMD for extended periods of time. The analysis must also look for reported instances of simulator sickness and attempt to correlate those reports with the control-display conditions present at the time of symptom onset. The analysis should use a sufficiently broad array of data collection methodologies so all relevant aspects of visual, auditory, cognitive, and haptic functionality are covered. Moreover, the analysis should specify measures that distinguish between device operability (i.e., is it as easy to use as practically possible) and user acceptance (Wickens, 1984). Using a mix of survey, interview, and observational measures, the analysis should canvas a number of different users who have had system experience to ensure that data on input, output, controls, displays, and processing deficiencies are identified.

Over the long term, it is reasonable to expect that VR will prove effective for training some tasks and not for others. As such, VR will not be the "silver bullet" that revolutionizes the military's approach to training nor radically alters how training is conducted at the squadron level. Consequently, an important goal of our AGSS human factors evaluation was to identify the *characteristics* of the user's tasks whose performance are aided, hindered, or unaffected by VR so the latter may be judiciously and cost-effectively applied to other VR training devices.

Objectives

Our main objectives were to: (a) determine any usability problems associated with the AGSS design; (b) identify skills thought to be most affected by use of the AGSS; and (c) ascertain whether AGSS implementation had any effect on flight-line training.

Methods

The bulk of the information obtained regarding the usability and skill "trainability" with the AGSS came from a Human Factors Survey. We also obtained some information on AGSS training effect (including some usability issues) through student surveys as they completed MQ AGSS simulator sessions, observations of these sessions, and grade folder evaluation.

Participants

Eleven experienced rotary-wing instructors volunteered to participate in the Human Factors Survey; five FEs and six AGs. Ten were crewmembers from the MH-53J weapon system and one was from the MH-60G. All survey participants had AGSS experience and extensive instructor experience in their respective positions.

Also, we surveyed 20 MQ students and 2 instructor gunners following each of the 5 MQ AGSS sessions which included 3 daytime and 2 nighttime rides. The students were primarily AGs, although a few pilots and FEs were also queried after AGSS familiarization rides.

Equipment

The AGSS (see Fig. 6) is designed to provide both basic skills and MR training to USAF AG and FE personnel. The AGSS is a reconfigurable (MH-53J and MH-60G) device with a three-degree-of-freedom (pitch, roll, and heave) motion base. The device simulates sound, recoil vibration, trajectory, and visual imagery for both the .50 caliber (cal) and 7.62 millimeter (mm) guns. Gun design mimics actual gun control, position, weight, feel, and operation. All ship, gun, and air sounds are heard by the trainees, including: engines, rotors, guns, landing gear, warnings, and airstream effects. The visual system provides both day and NVG night scenes including: outside view, inside cabin, gun barrel, fixed and moving targets, weapon effects, effects of hits or misses, and environmental conditions. The AGSS is designed to operate as a standalone trainer (AGSS only) and an integrated trainer (AGSS with a host cockpit).

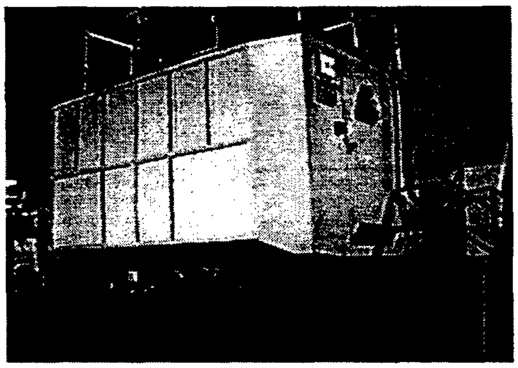


Figure 6. Outside View of the AGSS.

The main systems we examined in the surveys are presented in Figure 7. Although this conception does not include all of the systems that comprise the AGSS, it does cover those of greatest interest to the end-user, where the Trainee Station is central.

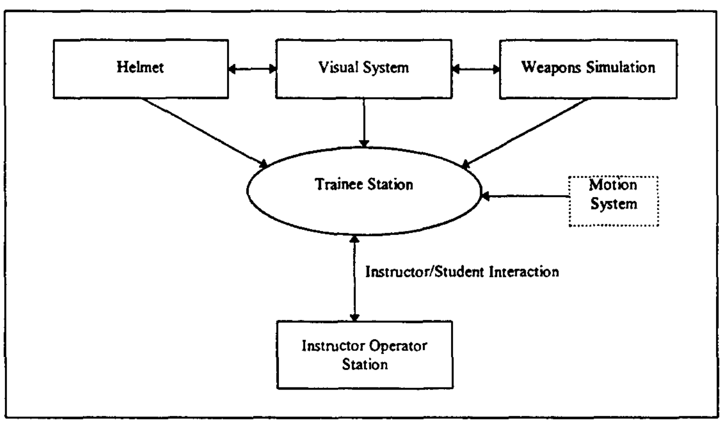


Figure 7. Main System Schematic of the AGSS (Silverman et al., 1996).

Surveys

The ***Human Factors Survey*** was designed to capture both quantitative and qualitative information about AGSS functionality and fidelity. The survey was divided into nine sections and questions asked in each section are briefly described below:

1. Background information *(questions regarding participant's age, current responsibilities, crew position, AGSS experience, flying experience, and visual history)*
2. Helmet *(1-5 Likert Scale questions on frequency of component breakage and malfunctions and comfort and fit problems)*
3. Trainee station *(1-5 Likert Scale questions regarding station components)*
4. IOS *(questions about the level of task saturation experienced by operators and 1-5 Likert Scale questions on the training effectiveness of specific components)*
5. Weapons simulation *(1-5 Likert Scale questions on the similarity of specific gun components to the aircraft for each gun type (.50 cal and 7.62 mm))*
6. Intructor/student interaction *(yes/no questions about instructor and student ability to communicate with each other)*
7. Visual system *(questions on physiological symptoms that occur as a result of flying in the AGSS and 1-5 Likert Scale questions on visual system elements)*
8. Training capability *(1-5 Likert Scale assessments on the training capability of the AGSS for 39 AG/S skills under both day and NVG night conditions)*
9. Miscellaneous *(questions on the overall feel of the AGSS and the motion system)*

For the ***Initial Impacts Assessment*** of the AGSS on MQ training, three separate surveys were used with slight variations depending on the scenario flown--two for the instructors and one for the students to complete.

The *Trainee Evaluation Questionnaire* included questions about AGSS system performance, instructor performance, and motion sickness, with space for general comments. The students were asked to rate components, instruction, etc. on a 5-point scale from "strongly disagree" to "strongly agree." The most useful sections concerned AGSS system performance and student comments.

Similarly, the *Instructor Evaluation Questionnaire* included questions about AGSS system performance, student performance, and AGSS training components (e.g., utility of the AGSS record-replay function). Since only two instructors actually trained with the AGSS during the survey time frame, the responses were of limited use.

The *Training Effectiveness Survey* was designed to capture instructor assessments of student performance on various AG/S skills performed in the AGSS from 1 (extensive assistance required) to 5 (no assistance required). Depending on the scenario, 24 to 44 individual AG/S skills were assessed. Identical surveys were designed for the flightline to determine learning curves and skill-level transfer of AGSS training to the flightline. However, problems with flightline scheduling and communication deterred the administration of the flightline questionnaires, resulting in an insufficient number of completed questionnaires to make meaningful assessments.

Procedure

The Human Factors Survey contained three types of quantitative responses: (a) yes/no responses; (b) never, rarely, sometimes, often, always-scale ratings; and (c) 1 (poor) to 5 (exceptional) Likert-scale ratings. In addition, each participant provided researchers with a substantial amount of qualitative information, including their informed opinions about device components or features, particular problems, and often suggested solutions. We conducted the survey using one-on-one guided interviews. The interviewer stressed the importance of making candid assessments and the fundamental goal of improving the AGSS and its use by students and instructors.

The Initial Impacts Assessment questionnaires were administered to students and instructors following their MQ AGSS training sessions. Analysis of MQ grade folders prior to AGSS implementation and after AGSS implementation also provided some information regarding AGSS potential skills to target with AGSS training in the future.

Data Analysis Strategy

We selected two statistical techniques to analyze our rating data. We used the Kolmogorov-Smirnov (KS), one-sample, goodness-of-fit test (Hays, 1973; Siegel, 1956) to analyze AGSS system component ratings. KS is a nonparametric test of the extent to which an observed distribution of scores deviates from some expected distribution. For each component, we used the KS analysis to test whether the obtained rating distribution was significantly different from a uniform rectangular distribution. This analysis let us determine whether observed ratings for a given survey item were significantly bunched toward the positive or negative end of the scale. Second, hierarchical cluster analysis was used to identify distinct skill groups based on SME ratings of training capability.

For the MQ student surveys and folder analyses, we explored some of the descriptive elements of the responses and folder entries, i.e., counts and sums. The results of these are described below as they relate to the major results of the Human Factors Survey.

Major Results

Table 1 summarizes the assessments of the five major AGSS systems as well as the device's overall Training Capability. Besides average overall ratings, which are generally acceptable, the table lists our general assessment (inferred by the researchers from participants' comments and ratings) and representative comments for each system.

Table 1. Summary Assessments for Each Major Human Factors Survey Section (Silverman et al., 1996).

System	Mean Rating (1 - 5)	General Assessment	Representative Comments
Helmet	4.3	Good (when working)	⇒ When properly fit, everything should be okay.
Trainee Stations	3.8	Good	⇒ Much better than first design.
Instructor Operator Station	3.9	Good (room for improvement)	⇒ Overall, it is nice. ⇒ Ownship flying with space ball is worthless.
Weapons Simulation	3.6	Needs Work	⇒ Not enough tracers per rounds fired. ⇒ Lead and lag problem.
Visual System	3.2	Faulty	⇒ Can't see targets until they are too close, even with instructor input.
Training Capability	3.5	Good (Qualified)	⇒ Primary impacts will be on crew coordination, scanner calls; target acquisition only if visual problems fixed.

Helmet

The helmet's overall rating was quite high, resulting in a general assessment of "good." However, low ratings for **essential** parts of the helmet, as well as participant criticisms, required us to include the "when working" caveat in our assessment. Critical components of the helmet, head tracker, and cathode-ray tubes (CRTs) are subject to frequent breakage and malfunctions. Moreover, the fitting and aligning processes are fallible, which can lead to additional helmet problems.

With regard to the head tracker, two main problems identified by participant comments and training observations were that it malfunctioned frequently during AGSS sessions and that it did not keep up with required scanning rates. In terms of the CRTs, the main problems noted were CRT performance deterioration over time, CRT failure rate, and user-perceptual disturbances that result from using an old and new CRT together.

Regarding helmet fit, the helmet pads, the occurrence of hot spots, the chin strap, and the electronic cables were problematic. The helmet pads were frequently cited in participant comments as being inadequate, with their method of attachment to the helmet singled out most often. Participants noted that the velcro tape comes off inside the helmet, and that the pads unglue and rip. Others reported that the pads are hard and uncomfortable. Hot spots were reported by survey participants when using the AGSS helmet even after short (e.g., 30 min) AGSS sessions. The helmet pads were again cited as a primary reason for the hot spots. In terms of the chin strap, participants proclaimed: "it's uncomfortable" and "it cuts into your skin." In addition, several participants reported problems with the amount of pressure they felt from the cables tugging at the backs of their heads when performing tasks in the AGSS.

We also asked participants to estimate the number of minutes required to put on and adjust the helmet during their first and subsequent encounters with the AGSS, and to indicate whether readjustments were necessary during training sessions. Average donning time for the first use was 11.1 min; for subsequent sessions, it was 3.2 min. The reduction in donning time was statistically significant ($t = 4.15$, $p < .01$, $df = 9$). Several participants reported the need to readjust the helmet during AGSS sessions. Three categories of adjustment problems appeared in participants' comments: alignment, the air pump, and familiarity.

Observations of training and student assessments of their experiences with particular device components from the Initial Impacts Assessment also indicated problems associated with the helmet. For example, 31 out of 40 surveys indicated "agree" or "strongly agree" to the question of whether the helmets required improvements, an average of 4.13.

Trainee Stations and IOS

The trainee stations received a fairly high overall average rating and participants' comments were consistent with this high rating. The majority of the comments focused on issues that could make the trainee stations *better*, rather than flaws about the design and functionality of the stations themselves, leading us to the unqualified assessment of "good."

The IOS had a slightly higher overall mean rating than the trainee stations. However, we assessed it as "good" with a "room for improvement" qualifier because participants gave unacceptable ratings to two critical features (the space ball for ownship flying and the second instructor position. In addition, although some of the comments were simple suggestions for *improving* the training capability of the IOS, many focused on changes that would *enable* training.

The most unanimously agreed-upon, poor feature of the IOS was the space ball for flying ownship during standalone AGSS training. It was singled out by the KS analysis as having significantly negative ratings. The central reason given for the overwhelming displeasure with the space ball for flying the ownship was the poor software modeling of aircraft aerodynamics which manifests itself in a tremendous difficulty maintaining proper flight control using the space ball.

There was also extensive agreement among participants on the lack of utility of the second instructor position at the IOS. The comments focused on the inability of instructors at this position to communicate directly with students in the AGSS. Echoing the thoughts of other displeased participants, one individual succinctly said, "You have the capability to hear, but you can't talk. This is dumb. Provide this capability."

Several questions were designed to tap into the existence of problems associated with locating the IOS outside the motion base of the AGSS (unlike other simulators at the 58 TRSS). Question topics included noise level, visual distractions, and the instructor task saturation due to the inability to see students or the requirement to monitor multiple screens for both student and aircraft performance. The only question having a large proportion of participants rating it as a problem was the level of instructor task saturation experienced at the IOS.

The mean levels of instructor task saturation experienced with one, two, and three students are shown in Figure 8. As can be seen, there is a linear relationship between the reported task saturation level and the number of students in the AGSS. Thus, as the number of students in the AGSS goes up, so does reported level of task saturation. Participant comments suggested that concurrent demands of flying ownship, laying down threats, and monitoring students can wreak havoc on the level of training that instructors are able to provide students in the AGSS. Figure 8 also shows the severity of reported task saturation increases when the one SME, who had a large role in AGSS IOS development and over eight years experience as a simulator instructor-operator, is removed from the calculations (i.e., the right bar graph of each pair in Fig. 8).

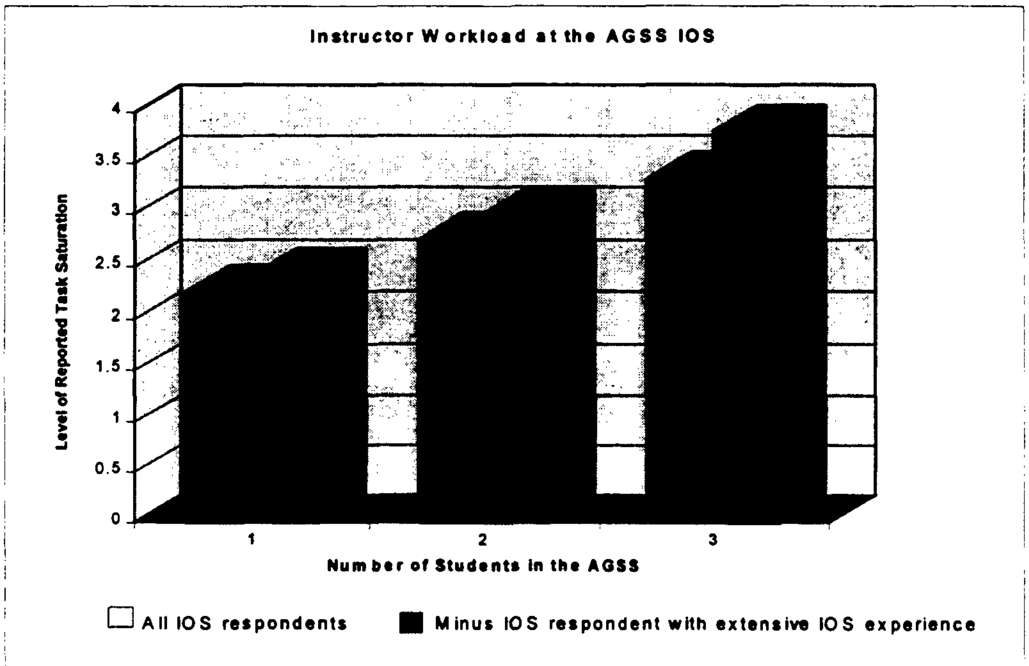


Figure 8. Levels of Instructor Task Saturation with One, Two, and Three Students in the AGSS (Silverman et al., 1996).

Relatedly, we also asked instructors several yes/no questions concerning <u>instructor/student interaction</u> (although not listed in Table 1). None of the participants reported significant problems with either hearing or understanding the students while at the IOS, nor were participants unable to hear or understand one another or the instructor while in the AGSS.

Weapons Simulation and Visual Systems

The overall "needs work" assessment of weapons simulation came primarily from the frequently cited modeling deficiencies of the bullet path. The fairly high overall average rating of weapons simulation, despite this rather major problem, stems, we think, from the fact that many of the gunnery-related training capabilities of the AGSS simply did not exist prior to its introduction.

Survey participant comments and the somewhat lower ratings led to an overall assessment of the visual system as "faulty." Participants noted the tremendous importance of visual input for many AG/S tasks, and given the current quality of the AGSS visual system, the level of training that could be provided by the AGSS for many AG/S tasks was limited. Visual resolution and object detection were two poorly rated features of the visual system.

Training Capability

Finally, overall training capability of the AGSS was assessed as "good" with qualifications. That is, participant comments and ratings reflected overwhelming agreement that the training *potential* of the AGSS was tremendous, especially in the areas (determined by hierarchical cluster analysis) of crew coordination, terminal area operation skills, and tactical skills (see Table 2). However, AGSS training potential could not be **fully** realized until improvements were made to the visual system.

Table 2. Highly Trainable Skills within the AGSS.

Crew Coordination Skills	Terminal Area Operation Skills	Tactical Skills
Voice Procedures	Taxi and Hover Calls	Gunner Cockpit Exchanges
Left/Right/Tail Calls	Go-Arounds	Defensive Countermeasures
Left/Right/Tail Interactions	Scanning	Tactical Knowledge
Crew Coordination: Overall	Tactical Approaches	Threat Breaks
		Radar
		AAA

In addition, the MQ Grade Folder Analysis revealed that from 96-04 through 97-01 (since the AGSS's implementation and major AG MQ curriculum changes), there have been multiple "unsatisfactory" rides. The primary skill areas that appeared to be problematic involved gun nomenclature and gun knowledge. However, other contributors were: crew coordination, calling approaches, missing calls, SA, confidence, and voice procedures. Many of these were indicated as highly trainable with the AGSS and could be targeted for future training in the AGSS.

Miscellaneous Items

The miscellaneous section of the Human Factors Survey contained several questions about the motion system. Six participants reported that the AGSS did not fly like the aircraft. Yet, eight participants reported being satisfied with the three-degree-of-freedom motion system. Combined, these two results led to the tentative conclusion that the limited motion base was not the primary reason why participants reported that the AGSS did not fly like the aircraft. Features that were cited included low vibration level, lack of gravity, and lack of seat-of-the-pants feel.

Self-reports of the frequency and severity of nine different simulator sickness symptoms (visual discomfort, headaches, double vision, blurred vision, disorientation, eyestrain or fatigue, neck strain, stomach discomfort, and nausea) during and after AGSS sessions were also obtained. To determine the number of subjects reporting symptoms, we first collapsed across all of the symptoms. We found that 54% of the survey participants reported symptoms sometimes or rarely (either during or after AGSS flights) and 46% reported never having symptoms This is consistent with the literature on simulator sickness, where the percentage of individuals reporting symptoms across a selection of simulators has been observed between 10% and 60% (Kennedy, Lilienthal, Berbaum, Baltzley, & McCauley, 1989).

When individual simulator sickness symptoms were scrutinized more closely, only two symptoms (eyestrain and visual discomfort) were reported by a significant proportion of the participants during AGSS training, and none were significant after AGSS training. Interestingly, eyestrain has been shown to be the most commonly reported simulator sickness symptom in past research (e.g., Gower & Fowlkes, 1989). Comments from survey participants primarily indicated that the AGSS visual system seemed to be the main cause for the reported eyestrain and discomfort. More research is recommended to determine which conditions (e.g., with or without motion) or tasks (e.g., scanning) are the most likely candidates to invoke simulator sickness symptoms in AGSS trainees, and what can be done to alleviate or minimize their occurrence. Notably, none of the participants in the Initial Assessment Impact during MQ reported any incidence of simulator sickness. The researcher who observed AGSS training sessions and demonstrations only witnessed one session in which the AGSS had to be stopped to accommodate an individual who was experiencing severe motion sickness symptoms.

Finally, student response to training in the AGSS, from both observations and Initial Assessment Impact survey analyses, were very positive. For example, 39 of 40 surveys indicated "agree" to "strongly agree" to the question, "The AGSS permits effective personalized instruction." And, 39 of 40 surveys also "agree" to "strongly agree" that "Without the dangers of live flight operations, effective instruction is promoted by the AGSS." Both instructors and students expressed some dissatisfaction with AGSS nighttime rides with the device in its current stand-alone mode. Instructors in particular questioned the value of interrupting student's nighttime schedules to fulfill a nighttime AGSS ride without an integrated cockpit.

In sum, the surveys and observations illustrate that: (a) most features of the AGSS are acceptable for training, but may benefit from improvement; (b) quite a few AGSS features are superb; and (c) the AGSS, overall, functions fairly well as an AG/S training medium.

Recommendations

Based on the above results, a number of AGSS hardware improvements having potentially high training payoffs were recommended. They are summarized in Table 3.

Table 3. Improvement Recommendations for AGSS Device Components (Silverman et al., 1996).

Component or Identified Problem	Recommendation
CRTs	Replace them frequently, ***and in pairs***, to prevent large visual discrepancies between the two eyes.
Head Trackers	Match technology (e.g., update rates available) with training requirements.
CRTs and Head Trackers	Monitor their performance to establish performance degradation time lines.
CRTs and Head Trackers	Provide users with an understanding of the limitations of the current technology so they don't have false expectations of the capabilities.
Cables	Option 1: Match technology (e.g., cable lengths available) with training requirements. Option 2: Provide a means of tethering the cables to prevent them from interfering with training tasks.
Helmet Pads	Invest in comfortable and high quality pads or try to incorporate VR technology into aircraft helmets.
Helmet Adjustment: Time	Be aware of the time involved for appropriate fit and alignment and its impact on scheduling and fulfilling training tasks.
Helmet Adjustment: Alignment	Establish standardized procedures and training to minimize adjustment time required and better convey concept to users.
Second Instructor Position's Lack of Communication Capability	Provide capability to communicate directly with students in the AGSS.
Problems with Space Ball for Ownship Flying	Option 1: Change software modeling of aerodynamics. Option 2: Change control input device to a joystick. Option 3: Do not use space ball for ownship flying.
High Levels of Instructor Reported Task Saturation	Option 1: Operate only in integrated mode. Option 2: Limit number of students in AGSS. Option 3: Provide two instructors for training. Option 4: Change layout of instructor screens.
Poor Modeling of Bullet's Path	Option 1: Change modeling of path. Option 2: Provide training on aircraft versus AGSS discrepancies.
Poor Object and Visual Scene Detail	Improve modeling and texturing of database and moving models.
Poor Visual Quality	Option 1: Improve visual system. Option 2: Improve other systems (e.g., CRTs).

In addition, several recommendations to AGSS *training* are provided below. These would enhance training, but do not require any modifications to current AGSS hardware. That is, observations of present AGSS training revealed that it is not necessarily geared toward the strengths of the device nor toward the current status of the device as a standalone trainer. The SOI was built with the AGSS's integration in mind, but this has not happened. Although integration and the above AGSS hardware modifications would offer increased training capability of the AGSS, they are not required to effectively train using the AGSS.

First, one should concentrate on developing standalone scenarios that have specific training objectives with a special emphasis on skills that the Human Factors Survey indicated were as highly trainable in the AGSS and those skills that seem to be causing problems on the flightline (grade folder evaluation data). The scenarios could be less "point-to-point" oriented (as we have observed) and shorter, perhaps repeating portions (i.e., better use of the record-replay function). An important and perhaps overlooked advantage of the AGSS in standalone mode versus the aircraft is that it allows for deliberate practice of critical AG/S skills.

Second, use of the AGSS for review/remediation rides should be expanded, perhaps in conjunction with the Helicopter Part-Task Trainer (HPTT) at the 58 SOW. Instructors can focus on AG/S skills that were a problem for the AG/S student without having to share time with other students. For example, one difficult AG/S skill to acquire is calling approaches. There is a current scenario that can be used that just has eight loops around an airfield and the AG is able to call in <u>all</u> approaches during the AGSS session (vs sharing time with FEs).

Third, the order of AGSS rides in the SOI should be changed; without integration, the current order of AGSS training rides does not make much sense. Instructors have suggested that AG/S need a lot of verbage and crew coordination training up front. Using the device more for this type of training early on in MQ would capitalize on the AGSS training strengths. We have observed that once the AGs are in the middle of aircraft night rides, coming back for night rides in the AGSS, especially without integration, has questionable training value.

Fourth, some aspects of gun malfunction training can be performed in conjunction with other skills performed in the AGSS. We observed one instructor talk through gun problems while AG student was scanning etc., even without the actual gun available. The instructor simulated the aircraft workload while talking through the gun malfunction procedures. The student responded quite favorably to the training following this AGSS session.

Implementing any one of these AGSS training modifications would enhance the training provided with the AGSS. Other potential outcomes include: more efficient and prescribed use of the AGSS, reductions in the number of aircraft sorties required for remediation, and stronger and closer links between the 58 TRSS and flightline training.

Technical Products

Two AGSS-relevant technical products are:

Silverman, D. R., Spiker, V. A., & Nullmeyer, R. T. (1996). *Human factors evaluation of the Aerial Gunner/Scanner Simulator*. (AL/HR-TR-1996-0146). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Silverman, D. R., & Spiker, V. A. (1997, September). A usability assessment of virtual reality simulation for Aerial Gunner Training. Paper presented at the *41st Annual Meeting of the Human Factors and Ergonomics Society*. Albuquerque, NM.

SPECIAL OPERATIONS FORCES NETWORK (SOFNET) TRAINING

As discussed in the introduction to this report, a high-fidelity networked simulation training capability has been developed by the 58 TRSS for its ART curriculum. This section summarizes the operational context within which this capability was developed, the study objectives, the environment in which SOFNET training was delivered, the study methods, and the survey results from participant crewmembers who participated in the SOFNET missions.

Background

Operational Context

In recent years, the ART curriculum at KAFB has undergone major shifts in both the substance and technology of training. Only a few years ago, ART consisted generally of academic classroom sessions covering aircraft-specific systems and EPs reviews. Following each half-day classroom session, students were required to apply these lesson topics, in a partial-task training approach, in their aircraft-specific simulator. This modular training method worked well for many years and the USAF units using this system were, for the most part, satisfied with the training provided. There was, however, a growing consensus in the small SOF community that the expense of bringing crews to KAFB each year for a repeat of the same curriculum was becoming prohibitive. The 58 SOW leadership responded by incorporating the unique network simulation capability to the ART curriculum.

The result has been a noticeable shift in philosophy from training aircraft systems to individual crewmembers to a CMT orientation for full-mission teams. Aircraft systems training was not deleted in this new training philosophy. Rather, systems malfunctions were now required to be exercised in real-time in the complex combat mission environment. Student crews performed corrective actions to instructor-induced systems malfunctions and EPs while other live combat stressors inflicted a toll on the mission team's actions.

Coincident with the new training philosophy, the 58 SOW was vigorously developing a sophisticated set of simulation and electronic training technologies. This has resulted in an unprecedented technological capability for ART that centers around five networked aircrew training devices: the MC-130P Combat Shadow WST and Satellite Navigation System (SNS), the MH-53J Pave Low WST, the TH-53A Operational Flight Trainer (OFT), the MH-60G Pave Hawk WST, and the HH-60G OFT. These devices include, among others, a full-fidelity simulation of aircraft and environmental systems such as Digital Radar Land Mass System (DRLMS), Night Vision Device (NVD), Infrared Detection Systems (IDS), Integrated Electronic Combat Simulation System (IECSS), and fully correlated, image-generator display systems.

The coupling of multiple, realistic training devices with high-fidelity, geospecific databases has enabled a unique CMT capability that combines constructive and live simulation of dissimilar devices on a common area of operations (AO) in real time and linked via a central Training Observation Center (TOC). With the integration of EW systems, countermeasures, improved visual systems, and correlated sensors, student crews could now perform real-world joint operations missions in a semi-immersion environment with a significantly increased training tempo. While it is assumed that increased capabilities in simulation training are critical to maximizing mission readiness, technology advances in networked simulation also offer promise for enabling services, commands, and weapons systems that are geographically dispersed to simultaneously train in realistic, virtual environments.

The combination of improved simulation technologies, a vision for implementation of distance learning methods, and a customer thrust for a mission-specific and joint operations training curriculum naturally led to implementation of the SOFNET training method as a proof-of-concept and demonstration capability. While these advancements hold great potential for training applications, little is known about how an established training program is affected when networked simulation methods are integrated into the curriculum. Accordingly, the 58 TRSS requested an impact assessment of SOFNET.

Research Issues

Given the recent advent of networked simulation technology, it is not surprising the existing research base is limited. Indeed, there have been few published accounts of the training effectiveness of such a technology, offering the researcher little guidance in developing measurement methods and analytic techniques. Recognizing this void, the Army and Air Force are nearing completion of a jointly sponsored project designed to assess the value-added to existing Service training offered by DIS as applied to the Close Air Support (CAS) mission (Mirabella, Sticha, & Morrison, 1997).

Their study queried participants following completion of a week-long, CAS training exercise. Focus of the assessment was not on performance per se but on users' perceptions of the training value afforded by participation in networked simulation and whether their training objectives were accomplished. Assessments of training value varied widely, depending on the participant's role. Of less interest than the specific findings are the lessons learned by the investigators. For example, they noted a mix of survey, observations, and interviews yielded the most valuable information; training value is perceived low when one user group's role is restricted to that of a "training aid" for another; and instruments need to be tailored to the service, training site, and mission of the user (Mirabella et al., 1997).

SOFNET Overview

In the context of this research base, Figure 9 depicts our SOFNET conceptual model. SOFNET training may be envisioned as the intersection of three modular components: Training Devices, Scenario Command and Control (C&C), and Mission elements. The first component, **Training Devices**, may be defined for our purpose as those participant crews who are represented as players in their respective training device(s) in the network environment. This critical component may be composed of any combination of network participants, and the training scenario is manipulated based upon the matrix of participants scheduled for the training.

It is reasonable to assume that the mission parameters and outcomes from the SOFNET training session will depend on the number and composition of participant crews.

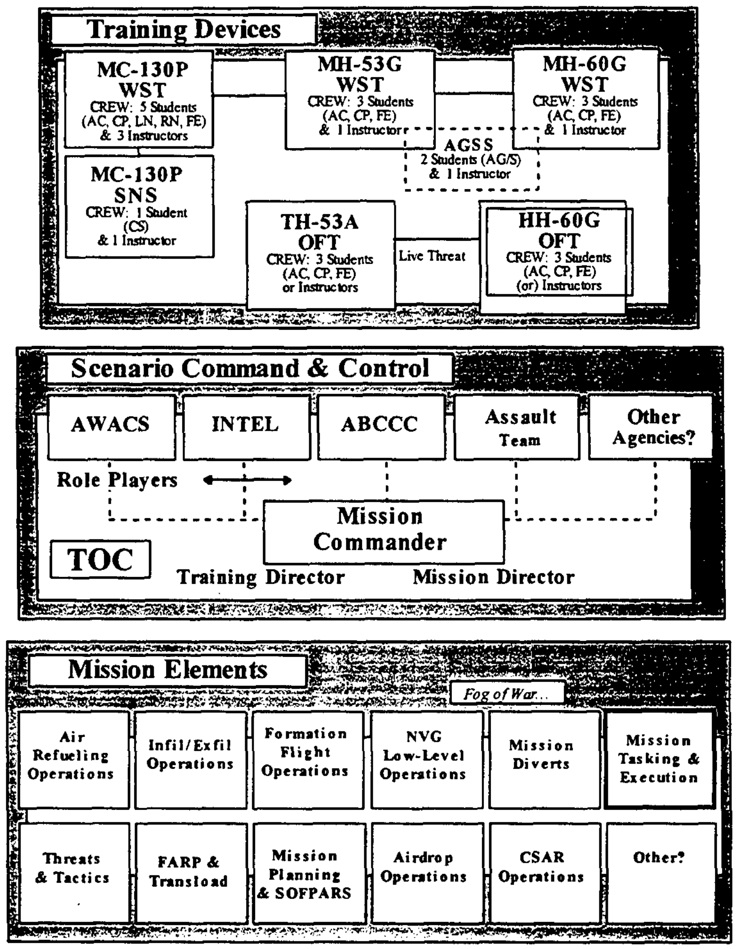


Figure 9. SOFNET Conceptual Model (Tourville et al., in press).

Next, the **C&C** component is envisioned as the various role players who are scripted into the particular combat mission training scenario, and either the Training Director (TD) or Mission Director (MD) who is tasked to weave these simulated operations coordination elements into a cohesive joint mission execution. Role players represent, for example, airborne command, control, and communication (ABCCC), a ground assault team commander, a simulated transload aircraft at a forward Landing Zone (LZ), INTEL, etc. These players are tasked to follow a fluid mission script as the training scenario develops, and intervene as necessary to steer the training mission to its intended conclusion. They must work as a team to integrate all SOFNET systems to achieve the desired training objectives.

Finally, the **Mission Elements** component may be envisioned as those particular mission objectives that may be accomplished, depending upon the skill and/or capability of the participants, or the C&C mission assignments in accordance with the mission tasking, and the development of the mission scenario as the script plays out in real-time execution. The C&C element may view this component as the "toolbox" of events that may be performed by a particular combination of participants in the network training. Specifically, the SOFNET training session may comprise any combination of aircrew participants with differing mission requirements and capabilities. The "toolbox" of mission elements that would emerge as possible joint training objectives might include, for example, AR, formation flight, threats and tactics, etc. Upon consideration of the matrix of available participants, the MD might consider tasking a different set of mission objectives for a particular training session.

To be effective, an evaluation methodology must support insertion of data collection "probes" into each component of the SOFNET training program. Not only must the data capture participants' reactions as they experience the different aspects of this novel training approach, it should also result in concrete suggestions for improving the training process that could be folded back into the curriculum in a timely manner. As with the joint service DIS study, a mix of interviews, observations, and surveys are needed to collect the user-reaction data.

Method

Mission Participants

Across the nine SOFNET sessions that were observed, a total of 99 crewmembers participated in the survey. This included 22 MH-53J pilots, 17 MH-53J FEs, 14 MH-60G pilots, 9 MH-60G FEs, 14 MC-130P pilots, 11 MC-130P navigators, 6 MC-130P FEs, and 6 MC-130P CSOs.

Mission Tasking

The scripted training mission was conducted in the Southwest USA database and entailed a high-level operation to conduct prisoner of war (POW) rescue operations in hostile territory. The scenario included forced entry into a POW compound to return coalition force detainees to friendly control. After recovery was completed, the assault force would transport recoverees and a medical team to an abandoned airfield for transload and evacuation. All forces were also tasked to be prepared to pickup downed crewmembers as required.

This tasking entailed a complex set of mission elements to be performed by participant crews. The MC-130P was to conduct NVG low-level operations to a precoordinated AR Control Point (ARCP), and provide refueling support for a two-helicopter formation (MH-53J and MH-60G) prior to their crossing into hostile enemy territory. After refueling, the MC-130P was to continue its NVG low-level operations to airdrop a reconnaissance-reception team at a transload airfield. This site is to be used by the helicopter formation, and others, to conduct blacked-out Infil landing operations and transfer of personnel to a waiting (simulated visual model) MC-130H Talon aircraft.

The MH-53J was to fly an NVG, low-level route as formation lead with the MH-60G as wingman. They were to fly to the preplanned ARCP, conduct refueling operations with the MC-130P, and "cross the fence" deep into hostile territory. This territory had a medium-level threat saturation, and formation tactics and altitudes were to be flown appropriate to the environment. After AR, their objective was to fly to the hostile POW compound and insert a special forces team via fastrope procedures. With helicopter air cover and fire support, the team would secure the compound and extract the friendlies back onto the helicopters. Following this, the helicopters would fly to the transload site for transfer and evacuation of the expatriated POWs.

The TH-53A or HH-60G OFT were represented in the visual model set as a threat helicopter (Hind) who is located at a precoordinated attack point. One of these devices was used to conduct live aggressor tactics and was manned by two instructor pilots and one instructor FE/simulator operator. Their objective was to run multiple attack runs on the helicopter formation, and attempt to divert them from their primary mission tasking.

Observations

Observations of the entire mission process--including briefing, planning, execution, and debrief--were made by the second author, an SME-researcher. He used a naturalistic observation technique to record salient data concerning the conduct and quality of the SOFNET training sessions. Most of the observations were made in the TOC, where student performance and mission progress could be perceived from the monitor displays, and from commentary by the MD and role player personnel. Additionally, MP and mission briefing sessions were observed for each of the student crews. These observations were made in the respective mission planning rooms, which presented the opportunity to openly discuss with students and instructors those training concerns associated with operational mission requirements.

Survey

Upon completing a SOFNET training session, the SME-researcher administered a 2-page questionnaire to each of the participating crewmembers. The first page of the questionnaire contained four questions regarding the briefing, planning, execution, and debriefing phases of the mission. Crewmembers were to record their comments, in free form, under each question. The questions were:

Question 1: How clear was your specific role as a network player in the mission script once the in-brief ended? What might have been done during the in-briefing to help clarify your role or requirements?

Question 2: Were there aspects of the mission script that required you to support the actions of another crew that were artificial or unrealistic?

Question 3: In what ways did your mission plan change as a result of interaction with other mission team players? What team interactions were helpful or not helpful? What other team interactions would have made the training more realistic?

Question 4: How is training in a joint networked simulation mode likely to affect your real-world chances for mission success? Which areas of the mission are most likely to be affected? Please explain.

Space was also provided at the bottom of the page for crewmembers to note additional comments concerning any aspect of the training session. On the second page of the questionnaire, crewmembers completed a 5-point rating scale that assessed their perceptions of the value of networked training relative to standalone training for 33 mission elements. A rating of "1" corresponded to "unacceptable;" "2" was "less value than standalone;" "3" was "same value as standalone;" "4" was "better value than standalone;" and "5" was anchored as "networking is essential for this element."

The 33 mission elements covered all aspects of the mission. These included major mission functions (e.g., mission planning, tactics planning, LL planning, multiship tactics); tactical tasks (e.g., AD operations, AR operations, formation flight, systems malfunction); and higher level cognitive processes (e.g., SA, crew coordination, TM, mission team coordination). Next to each rating, space was provided for crews to note any comments that might help explain or amplify their judgments.

Major Results

Quantitative Rating Data

The participant survey results showed that *all* of the mission elements received high ratings, with all but two of the mission elements rated positive (i.e., above the scale midpoint, 3.0) by the SOFNET crews. Table 4 lists the mission elements in descending order of mean rating. Since some of the elements were not applicable to certain crew positions and weapon systems, the total N is typically less than 99.

Only AD Operations and Checklist Procedures failed to achieve a statistically significant positive rating. Statistical assessment was based on a simple t-test, in which a Bonferroni adjustment was imposed to control for multiple testing (Harris, 1994). Importantly, overall value of networking simulation was rated positively. This assessment was particularly conservative, since the maximum rating scale for this element was "4." This was done because we felt the "5" rating used for the other elements, signifying "networking is essential for this element," was not a logical possibility for an overall value assessment. Despite this conservatism, a statistically significant positive rating for the overall value was still obtained. Hence, it is clear that most participants felt that networked CMT represents a substantial benefit over traditional standalone simulator training.

The most positively rated elements involved coordination among multiple players. The seven elements with mean ratings of 4.0 or higher were Multiship Tactics, AR Operations, CSAR Operations, Formation Flight, SA, C&C, and Mission Team Coordination. All of these elements require integration of crew efforts to be successfully completed and, as such, their positive ratings are consistent with the stated purpose of the network training to support integrated crew operations.

Table 4. Mean Rating of SOFNET Mission Elements (Tourville et al., in press).

Mission Element	Mean Rating (1 = lo, 5 = hi)	N
Multiship Tactics	4.2	84
AR Operations	4.1	93
CSAR Operations	4.0	59
Formation Flight	4.0	74
Situation Awareness	4.0	97
Command & Control	4.0	95
Mission Team Coordination	4.0	93
Transload Operations	3.9	61
Time Management	3.9	97
Mission Debrief	3.9	93
Mission Briefing	3.8	96
Secure Comms	3.8	84
In-Flight Formation	3.8	96
Crew Coordination	3.8	96
Mission Diverts	3.8	77
Mission Planning	3.7	95
Threat ID and Response	3.7	87
Infil/Exfil Operations	3.7	74
Threat Avoidance/Evasion	3.6	93
Fuel Management	3.6	90
Low Altitude Operations	3.5	93
Weapons Employment	3.5	65
Tactics Planning	3.4	94
Night Operations	3.4	92
Radar/FLIR Interpretation	3.4	85
Chaff/Flare Management	3.4	82
Low Level Planning	3.3	93
Airdrop Operations	3.3	36
Terrain Familiarization	3.3	90
Systems Malfunctions	3.3	95
Minimum Wx Operations	3.3	81
Checklist Procedures	3.1	94
Overall Value	3.7	99

Although the ratings suggest that the SOFNET training was highly valued overall, we probed the survey rating data further to ascertain if there were any notable differences in ratings across crews, weapon systems, and crew positions. With regard to crew, the average rating of the 32 mission elements varied only between 4.1 and 3.5 across all nine crews, where there was no clear evidence for either a downward or upward trend over time in the ratings over sessions. Similarly, there was no statistical evidence of any differences in perceived training value across weapon systems. This was somewhat surprising, since we had suspected that MH-53J crews might have given SOFNET a higher rating since the mission scenario was originally designed to enhance the Pave Low's training capability.

However, when the survey data were broken out by weapon system *and* crew position, we found that pilots exhibited a strong preference for integrated over standalone simulation training (M = 3.9) compared to their nonpilot counterparts (M = 3.6). This difference is particularly evident for the MH-53J and MC-130P weapon systems.

Qualitative Comments on the Mission Elements

Fully one-third of the 99 participants took the opportunity to amplify their ratings of the mission elements with comments. Commonality analysis showed that crew responses could be grouped into five higher level categories: planning and briefings, visuals and terrain, tactics and operations, communications and procedures, and team interactions.

Examples of the most frequent comments included a desire for more planning time and a dissatisfaction with computerized planning products. There was a mixed reaction to the quality of the visuals, with a number of crewmembers expressing particular difficulty seeing the terrain while flying low level and performing formation flight. Many of the crews indicated a desire for a more intensive threat environment. While comments regarding the difficulty in achieving reliable communications were reported, many noted that this provided an unintentional replication of the "fog of war." Finally, despite many problems encountered during the simulator sessions, most participants thought the networked environment provided an excellent opportunity to practice team interactions, thereby enhancing team awareness and offering opportunities to try out different solutions to problems that arose during the scenario.

Qualitative Responses to Networked Training Questions

Crews were quite responsive to the comment portion of the survey, with 95% of the available slots (4 questions x 99 participants) containing comments. The following paragraphs are intended to convey the essence of some of the major points that were reported.

Clarity of mission script. For the most part, all SOFNET crews indicated that their specific role was clear from the briefing. A typical response was that it was "mostly clear except for . . . " The specific response varied across crews and included issues such as "no h-hour established," "comm package was weak," "limited imagery in the objective area," and "an Air Tasking Order should have been included in the briefing package."

Did mission script require unrealistic actions? An impressive two-thirds of the participants responded that the mission script did not require them to perform in an artificial way to support actions of crews in the other WSTs. This suggests that the SOFNET mission script struck a fairly even balance between establishing objectives for an actual combat mission while ensuring that each participating weapon system had some meaningful tactical tasks to perform. Critiques were noted by the other one-third of the participants, and they fell into one of five categories: simulator capabilities, mission conditions, terminal area tactics, role-playing, and mission objectives.

With regard to *simulator capabilities*, five of the participants noted that although mission demands required the MH-60G have the more capable 701C engine (which is in the actual aircraft), the WST has the less powerful 700 engine in the aero package. The result was an inability to perform the required AR on a consistent basis. Several crewmembers noted the distraction caused by hearing "admin traffic" between the instructors in the different WSTs. Another criticism was that the MC-130P, MH-53J, and MH-60G WSTs "did not have the airspeed comparisons that match the aircraft." This mismatch required the MC-130P to fly close to stall speed while performing ARs with the helicopters.

With regard to *mission conditions*, four participants cited the unrealism of the original temperature and density altitude parameters within the LZ given the limited power available for the MH-60G. One helicopter pilot indicated they had to "lower the temperature" to perform the required pick-up. Another critique concerned the MC-130P's requirement to penetrate a high threat area to perform the AD. While several participants cited the unrealism of this requirement, they also noted that it was good training and "lots of fun" given the briefed threat scenario.

Within the *terminal area*, several tactical aspects of the mission were criticized. These included reports of "bogus" indications from one of the surface-to-air missile (SAM) sites, the use of secure radios during calls for fire support, and reports that the simulated threat level in the DZ was too high for some weapon systems' tactical requirement.

Several aspects of SOFNET *role playing* were critiqued. One crew experienced problems with their visual system, which forced them to "role play" more of the scanners' functions. Also, the lack of wraparound visuals in the WST gave them less scanning capability than they would have in the aircraft. Several MC-130P pilots noted it was unrealistic to drop the SF team to secure the forward area refueling point/transload site given the threat level in the area.

With regard to the SOFNET *mission objectives,* several MC-130P pilots noted that one of the required helo ARs was unrealistic given the threat level in the DZ. Another MC-130P pilot complained about the lack of a role player in the scenario whose inputs would "affect our (MC-130P) decision making realism." In addition, several pilots doubted that both the MH-60G and MH-53J would be able to fit into the terminal area to drop the special forces team in the building.

How plan changed as a result of interaction with other players. The vast majority of the survey participants responded to one or more parts of this question. With regard to mission plan changes prompted by team interactions, four changes were cited most often, all of which were considered positive from a training impact. The specific change to the plan varied across crews, reflecting minor modifications to the mission script imposed by the instructors.

First, many crewmembers lauded the training benefits associated with having to slip the MC-130P's AR control time because of the late takeoff time by the helos. Second, the nature of the AR (i.e., orbit point, track) had to be changed due to the helo "going single engine." Third, one helicopter crew noted that they had to change their planned ingress after diverting to pick up a downed Hind crew. Fourth, another crew had to alter its plan in-flight in order to perform an unplanned AR. In all cases, these changes to the mission plan were viewed favorably since they are "very helpful for coordinating with other aircraft."

Other frequently lauded aspects of networked training included ability to meet face to face with crews from other weapon systems, training benefits from the manned aggressor impact on ingress route and enroute times, and opportunity to interact with role players.

The most frequently cited critiques included difficulties of communication with some crew positions, the need for more planning time to iron out weapon system differences, the likely benefits of having a mass briefing, and the need for other role-players to be included in the scenario, such as ABCCC and AWACS.

How training in joint networked simulation mode affects real-world chances of success. All but four of the survey participants responded to this question. Regarding the first part of the questio.., respondents were uniformly positive regarding networked simulation impact on perceived chances for mission success. In fact, only 4 of the 95 participants responding to this question expressed negative sentiments concerning impact of networked training.

SME Observations

Despite some inevitable technical problems during the early stages of SOFNET implementation, there were a number of areas that clear improvements, indicative of a true "learning curve," were exhibited. For example, over the course of the study, the MD continually worked to enhance the mission in-brief. At the beginning of the study, the in-brief consisted of paper notes "read" to the crews. However, by the end of the period, the in-brief had progressed to high-quality Powerpoint slides with embedded cartographic and objective area photographs.

Another area of improvement concerned the quality of MP materials. Early on, SOFNET instructors were dissatisfied with the limited materials available to support real-world-type planning activities. Once network problems were isolated and corrected, the MD, with assistance and input from several instructors, began to improve MP materials with an enhanced set of crew-tailored mission packets that included weather, route details, intelligence information, etc.

Soon after the study began, it became evident that the same network problems were being encountered on a session-to-session basis, and these problems were simply not getting fixed (e.g., spontaneous system crashes, unrecoverable loss of communications channels). At that time, the MD requested a dedicated technician sit in the TOC for the duration of each SOFNET session to observe, in real time, the problems being encountered. This individual also sat through the WST-specific debriefs at the conclusion of training. Through such continued contact, where lessons learned were shared among all participants, the network failure rate was significantly reduced.

A fourth area of improvement involved instructor communications. Early on, the MD established a one-hour meeting--prior to the crews' arrival at the simulator--that was attended by the MD, all SOFNET instructors, and network technicians. This meeting served multiple purposes, including the provision of additional instructor training on the operating idiosyncrasies of SOFNET, identification of current system problems, and coordination of training strategies concerning joint mission operations.

Implications for Training

Areas Where Networked Training is Beneficial

Taken together, crew comments, mission element ratings, and observations provide a fairly clear indication of areas where networked training is beneficial for CMT. As noted by the ratings, the main areas where networked training value is high include multiship tactics, AR operations, CSAR operations, formation flight, SA, C&C, and mission team coordination.

From the comments, other benefits of networked training include its capability to simulate the "fog of war," promote the practice and honing of key tactical skills, improve "comm planning" and instill "flexibility in C&C interactions, learn the capabilities and limitations of dissimilar platforms, increase opportunities to do contingency or "what-if" planning, coordinate higher level tactics among crews and joint commands, test new tactics, and promote improved methods of task prioritization.

Areas Where Networked Training Needs Improvement

Despite the positive benefits lauded by participants, our observations, coupled with crew comments, revealed a number of problem areas that need to be addressed in future applications of the networked simulation technology. Presently, SOFNET mission training practice is not focused on any set of tangible training objectives, nor is it geared to take full advantage of the strengths of the training methodology. SOFNET training also requires a great amount of time to develop and conduct (the training day is often greater than 10 hours). The full-day experience of SOFNET training precludes an opportunity to provide further focused training on, for example, aircraft systems and EPs. Also, the SOFNET scenario is currently scripted in an unclassified database using a fictitious country setting with made-up names. Crews reported that fictional names simply make the scenario more difficult to remember and harder to relate to any personally held, "mental model" of the world. Conversely, students expressed a preference for conducting complex mission operations in real-world "hot-spot" locations with which they are familiar.

The ability to encounter threats and apply tactics in a real-time, networked environment is a simulation training advantage that cannot be replicated in the aircraft. However, many crewmembers report that the mission script should include a greater threat saturation level for all network players. The threat modeling capabilities of the network simulation offer a significant advantage over standalone training because live threat encounters require students to think and act in real-time to the unknown attributes of combat adversaries--a characteristic that can only be partially replicated by using a moving model system. Finally, it has been reported that the present SOFNET mission scenario is "fun," but is only challenging for pilot students. Some non-pilot students are not receiving effective training, and are often only "along for the ride."

Recommendations for Improving Delivery of Networked Training

We have established that the SOFNET training method is an effective means for mission crews to train for joint operations coordination and procedures. Certain changes are necessary, however, before the full benefit can begin to be realized. For example, the mission training scenario should be restructured to include a greater amount of systems and EP training. Also, a greater emphasis should be placed on development of explicit training objectives and focused strategies.

Additionally, student participation roles should also be equalized across WSTs, such that all participants are involved and tasked to perform at similar training levels. Actual time spent in the simulators performing the mission scenario should be shortened. It is counterproductive to prepare students to perform, for example, in the threat environment, and then task them to fly for two hours in the simulator before anything interesting happens. Finally, there is a great need for more face-to-face time between mission crews as they plan and prepare for their joint mission operation. The actual time cut from the simulator period would be better utilized by providing a longer, and more focused, team MP session.

The 58 SOW is presently conducting an in-depth assessment of its training requirements to determine the optimal methods for enhancing the impact of networked training. It is clear that the 58 SOW is "out in front" with regard to the implementation of advanced networked training. Indeed, the Wing owns a one-of-a-kind training capability, that with relatively minor adjustment, can become a model for establishing advanced Distributed Mission Training (DMT) foundations and principles. This training has the potential to not only significantly improve combat mission readiness of the warfighter, but also increase effectiveness of joint combat mission operations.

Technical Products

The following two products were delivered during this project period:

Tourville, S.J., Spiker, V.A., & Nullmeyer, R.T. (in press). *Analysis of the Special Operations Forces network training for joint mission operations simulator training.* (AL/HR-TR-1997-xxxx). Mesa, AZ: Aircrew Training Research Division, Armstrong Laboratory.

Spiker, V.A., Tourville, S.J., & Nullmeyer, R.T. (1997, December). Networked simulation and combat mission training. *Proceedings of the 19th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

RESEARCH IMPLICATIONS

This closing section discusses prospects for conducting future research activities at the 58 SOW, focusing principally on empirical data collection efforts based on the work described above. We begin by discussing four issues that transcend individual areas of inquiry. These are: (a) reviewing training requirements and performance objectives, (b) integrating ground training and flightline operations, (c) improving training effectiveness information, and (d) developing strategies for effective utilization of existing training tools. Following that discussion, we then describe four potentially high-payoff R&D activities that we expect to pursue with the 58 SOW and which apply findings from the studies described in this report. These planned research activities include (a) extending the T^2RM methodology to rotary-wing aircraft, (b) revising the MC-130P ART curriculum and assessing the impact, (c) developing measures and metrics to assess training effectiveness, and (d) developing an advanced MP curriculum.

Recurrent Themes

Crew and Team Training Objectives and Student Performance Criteria

The MC-130P CRM/CMT crew performance study and the SOFNET survey dealt with crew-level and multicrew training, respectively. Each study focused on a particular day of training for mission-ready crews in the week-long ART syllabus. As part of this research, we contacted training managers, instructors, and courseware developers to obtain formal training requirements for these portions of ART. We also sought performance criteria against which crewmember behaviors could be compared. We found a formal requirement for CRM training, which is satisfied by a half-day block of academic instruction at the beginning of ART. However, we were unable to document formal CRM training requirements or performance criteria associated with the simulator-based training that we studied. Instead, it appears that the driving principle was to use simulation to provide beneficial experiences for refresher crews that could not be provided at their home units.

This pattern is consistent with a recent description by Helmreich and Foushee (1993) of CRM training across the aviation industry. They stated that, while there has been a great proliferation of CRM courses, there has not been a parallel growth in the use of flight simulators for CRM practice and reinforcement. Nevertheless, they applauded the emergence of a new generation of CRM training that: (a) integrates CRM academic instruction with the use of simulation, (b) is more specific in focus, and (v) addresses optimum behaviors (e.g., behavioral markers). An additional characteristic of this new generation of training is the removal of distinctions among technical training, evaluation, and CRM instruction (FAA, 1996).

We see enormous potential benefit in following a similar path in military flight training. We contend that developing specific training objectives and specifying desired student performance criteria are essential elements of effective simulator training. The value of such specification is not limited to CRM training, but rather, should have broad applicability. For example, the AGSS human factors evaluation revealed a consensus among instructors that the AGSS has great potential to effectively train crew coordination, terminal area operations, and selected tactical skills. This premise will be translated into training benefits only to the extent that these objectives are reflected in the AGSS syllabus of instruction.

Integration of Ground-Based Instruction and Flightline Operations

The flying squadrons within the 58 SOW can be viewed as immediate customers for much of the ground training provided by the 58 TRSS. Trouble spots on the flightline are obvious places to review ground training objectives, behavioral criteria, and instructional practices. For example, a recent review of student folders reveals multiple "unsatisfactory" rides in the flightline phase of MQ training that can be attributed to a lack of crew coordination, faulty mission calls, poor SA, lack of confidence, and poor voice procedures.

The AGSS survey results suggest that many of these areas are correctable using simulation training. The next logical step is to review training objectives to see if, at least on paper, there are ground-based training objectives to address the training needs. Starting with stated training objectives, a training system trouble-shooting algorithm may reveal multiple places to change ground training to better prepare students for the flightline.

At a more general level, the Model Aircrew Training System (MATS) program produced a blueprint for integrating flight simulation into formal school aircrew training (Fishburne, Williams, Chatt, & Spears, 1987). The general goal was to create a structure for students that would provide a context for the training events that make up the simulator phase of training. The outcome would be an increased meaningfulness of the training, which should in turn facilitate retention of skills. A general review of the sequence of training events should therefore be conducted, focusing on those areas that are repeating problems on the flightline.

Specification of Student Performance and Training Effectiveness Information

In the two research projects (CRM/CMT and SOFNET) that encompassed crews undergoing ART, crew performance is not evaluated and crew knowledge is not tested. There is no pass or fail contingency for this stage of training. Rather, the payoff for both the home unit and the crewmember is to log the accomplishment of selected training events, many of which cannot be accomplished in the home unit due to the lack of flight simulators in the field. This event-driven approach to ART is a product of both training regulations and tradition.

The diversity and richness of the empirical results from the CRM/CMT study suggest that data reflecting crew performance in simulated tactical mission environments can be a valuable resource for training management. The five T^2RM subprocesses and associated T-MOT behaviors addressed here are only a first attempt to quantify the performance of the crews undergoing crew or team training. Even at this early stage of data generation, it appears that results can be fed back to the training community to modify CRM training content. For example, we identified a number of effective crew characteristics not exhibited by most crews now, and presently not taught to crews. The AC's use of verbal backups, the LN integrating inputs from the RN and CP, and the CSO "passing comm" in short bursts are but three of many insights captured in T-MOT observation protocols that could be included in MQ training as techniques to enhance a crew's tactical performance. Identification of such CRM behaviors would clearly arise incrementally, and would continue to contribute content to CRM training at the MQ level.

A second area of interest was post-mission debriefings. In both CRM/CMT and SOFNET training, there appeared to be a lack of consensus among instructors concerning CRM performance levels that should be exhibited by crewmembers, and whether performance feedback should be provided at the end of the training scenario, especially with respect to CRM behaviors. The modal instructor response was to not deal with CRM issues as part of the debriefing. A probable outcome of this decision is that exposure to these ART scenarios will have little impact on subsequent crew performance back at their operational units.

Enhanced student performance criteria and data capture could serve at least two vital roles within the training organization. The first is to provide performance feedback to crews at the end of training scenarios. Such feedback is probably an essential element of effective simulation-based tactical training. Second, such data should provide valuable feedback concerning training system performance to support training system management and decision making. As training budgets continue to shrink, such data can either be used to (a) justify continued investment in a given arena, if in reality there is a clear need for and benefit to training; or (b) identify areas of less need which can thus be reduced to free up resources for areas of higher need and payoff. The labor-intensive nature of SOFNET training (approximately 10 instructors and role players plus another dozen or so students) makes it a particularly important area for increased insights concerning training needs and benefits.

Identification of Training Strategies

Each of the three research studies discussed in this report produced recommendations to the sponsoring organization on how to increase the effectiveness of existing training technology through modified training strategies. CRM represents the most mature area investigated, and even this area is still undergoing fundamental changes in pursuit of quantifiable training effectiveness (Helmreich & Foushee 1993). CMT, multicrew training, and the use of VR technology in aircrew training are leading edge activities for actual training organizations, and as such, have even less developed instructional strategies to guide their use.

Our recommendations for CRM training involved both general process changes and more specific training content changes. The entire area of multiship training is in an embryonic stage of development. Many CRM training issues are still unresolved and are compounded in SOFNET by the added complexities of multiple crews. For both CRM and SOFNET training, we recommended a policy clarification by the Air Force to make crew or team coordination a legitimate objective of the simulator-based training experiences in ART, and then train instructors on how to be CRM debriefing facilitators.

Our AGSS recommendations entail taking advantage of current capabilities for crew coordination, terminal area operations, and tactical skills training while attempting to improve the system's visual resolution to expand the tasks that can be trained in the device. A major issue to be addressed is to determine those tasks best trained with the AGSS in a standalone mode versus those best trained when the device is integrated with other crewmembers in an electronically linked simulator. As visual resolution improves, another issue becomes whether improvements are sufficient to expand training into areas such as target acquisition and target identification.

We anticipate that development and validation of SOFNET instructional strategies will be a fertile area of research for the foreseeable future. In addition to team training issues involving multiple crews, the training and research communities need to consider the inclusion of other combat team members such as Intelligence and customers from other services. The SOFNET environment only partially captures the information flow via role-playing instructors. Yet we know that such information exchange plays a major role in the crews's decision-making processes. To gain a broader perspective of such combat teams and their dynamics, we will eventually need to consider training processes in light of distributed interactive simulation (DIS).

Follow-on R&D Activities

Revise MC-130P CRM Training in Mission Qualification

Analyses of MC-130P crewmember performance during a simulation-based CMT scenario produced results that supported some aspects of current CRM training, but also brought a number of current practices into question. Specifically, a strong correlation between overall CRM process ratings and the total mission performance rating suggested that there was indeed a value of improved CRM behaviors to mission performance. As we looked at these effects in more detail, crew-level analyses revealed that four of our five CRM subprocesses--SA, TM, TE, and FA--were highly predictive of mission performance, but C3 was not. Of these predictive processes, the dominant elements varied to some extent across mission phases (MP, AR, LL, AD, and I/E). Crew position analyses revealed that C3 *was* important for particular pairs of crewmembers at specific times in the mission. The CRM behaviors that led to high or low ratings were compiled, and the vast majority were highly crew position- and mission element-specific. Instructor reviews of these results indicated that the patterns in our data reflect operational reality.

In MQ training, CRM instruction is presently based on an AFSOC CRM workbook. The same workbook is used for all weapon systems and crew positions. Comparing our results to the topic areas covered in the workbook, we found only limited overlap. Consistent with our findings, the workbook addresses SA. However, TE and TM, which were two powerful predictors of mission performance, are not addressed. Communication, which we found to be a neutral factor at a general level, was a major topic in the workbook. Unfortunately, the important pairs of crewmembers and critical mission events were not covered at all. Similarly, the FA behaviors that distinguished the most effective crews were not covered in this general AFSOC overview. The workbook did provide a generic discussion of interpersonal dynamics and problem solving--two areas not addressed in our study.

The 58 TRSS has recently established a Working Group to review the content of CRM training in MQ in light of our findings. It is our belief that there is a great deal of specific, CRM-relevant information that can help crewmembers be more effective. Most of this information is weapon system- and mission-specific, and much is crew position-specific. We anticipate that student critiques will improve, and the number of problems on the flightline attributed to poor crew coordination will decrease if this information is incorporated into a revised CRM curriculum. AFSOC/DOT has reviewed and approved the basic changes we envision, has pledged its support of a tryout using MC-130P crews, and has requested a broader application of basic CRM training principles to crew training for other AFSOC weapon systems.

Extend the T^2RM Methodology to Rotary-Wing Aircraft

Having demonstrated a strong relationship between T^2RM processes and mission performance in the MC-130P, an important next step is to show that the process-performance relationship we observed is not limited to a given weapon system, scenario, or training regime. For example, the six-crew complement in the MC-130P WST is twice that of the MH-53J WST. Do team process effects still exist when crew size is reduced by half? Future work is needed to delineate the boundaries on the size of the T^2RM process effects that one can expect to find in these other contexts.

For our purposes, we have chosen to look for a strong process-performance relationship in rotary-wing crews, and more specifically, the highly tasked MH-53J Pave Low weapon system. To date, virtually all prior CRM research has focused on the process behaviors of fixed wing crews, with few studies devoted to their rotary-wing counterparts, either civil or military (Spiker et al., 1996).

In a previous analysis of MH-53J refresher training, we noted that crew coordination played a significant role in the mission-training success of the crews (Silverman, 1994). However, the content of the CRM course was not specifically tailored for rotary-wing needs, having been borrowed wholesale from (the mostly) commercial, fixed-wing community. A detailed front-end study, conducted along the lines of our work for the MC-130P, is needed to identify the T^2RM subprocesses that are most important for rotary-wing operations. When coupled with an extensive empirical data collection effort, results should yield valuable insights regarding optimal strategies for improved training and reinforcement of the associated T^2RM behaviors.

From a research standpoint, the study of T^2RM processes in rotary-wing applications poses a number of interesting conceptual and theoretical questions. For example, the crew complement in the MC-130P WST, six, is rather large. Do team process effects still exist when the crew size is reduced to three, as is in the MH-53J WST? Are crew position effects more pronounced with smaller crews? Does the nature of fixed-wing operations allow more compartmentalization of crew position duties than might be found with helicopter operations?

On the surface, there is no a priori reason why the MC-130P study results would be characterized by an unusually large process-performance association or by predominance of any one crewmember. In fact, this relationship and importance of individual crewmembers might be even stronger when crew size is smaller, as each crewmember's role is proportionally more vital. Group size effects could perhaps be estimated by comparing the MC-130P study results to a comparable one with MH-53J crews. Future work is needed to delineate variations in the magnitude of T^2RM process and crew position effects in other contexts.

Develop Better Information to Assess Training Effectiveness and Manage Training

Prior AFRL/HEA aircrew training systems studies showed that one of the most pressing requirements for gauging effectiveness of any training intervention is to have some sort of integrated information management system (I^2MS) in place at the wing or squadron level (Bruce. 1989; Bruce et al., 1991; Nullmeyer, Bruce, & Rockway, 1991). Failure to implement such systems throughout the USAF training community can be linked to the *absence* of four assets:

1. data that reliably measure a crewmember's task proficiency at a given stage of training,
2. the resources to analyze those data,
3. a centralized system for using the analyzed proficiency data to evaluate the overall effectiveness of the training intervention, and
4. a plan for circulating evaluation results to decision makers and stakeholders so that training delivery may be improved.

As documented in Bruce (1989) and Bruce et al. (1991), there is a lack of proficiency-based "*data probes*" that permit a reliable and valid assessment of any training intervention effectiveness. There are multiple reasons for this deficit, including: a dearth of process-based, aircrew combat capability measures; treating completion of a given training event as a fixed requirement rather than as a dependent measure of skill; overreliance on qualification checkrides as the primary means of assessing combat skills; and an orientation toward ensuring adherence to minimum performance standards (Bruce, 1989). To be useful, an I^2MS must possess the proficiency data needed to establish the skill levels of the aircrews. In this regard, Bruce (1993, personal communication, August 24, 1994) has noted that "there is a major disconnect between the stated objectives of the training program, producing combat capable aircrews, and the means to assess these capabilities."

Even if such proficiency-based skill data are obtained, units typically lack the *resources* to analyze the data. Bruce et al. (1989) noted that some commands have developed trend analysis programs that measure and monitor deviations from acceptable training standards over time. While the framework for the analyses is logically sound, many operating units lack personnel, expertise, and computer support to consistently analyze such trends or to circulate reports concerning their statistical status (P. Bruce, personal communication, August 24, 1994). Indeed, given the amount of paperwork needed just to document that a specific training event occurred as planned, flightline and training personnel are truly "awash in a sea of requirements."

Besides resource constraints, a perennial stumbling block to the evaluation of training system effectiveness has been lack of an integrated or *consolidated evaluation function* within the unit (Spiker & Nullmeyer, 1995a). By no means unique to the military, failure to elevate evaluation to the same level as acquisition and resource management has hindered the development of a centralized function that can be used to evaluate a training systems's merit and worth (Nullmeyer, Bruce, & Rockway, 1991). Presently, evaluation functions are fragmented in most units, with a corresponding diffusion of responsibilities. Even if a centralized database were developed, a centralized office for processing, analyzing, and disseminating evaluation results would be needed (Bruce, 1989). Such an office would require technical evaluation expertise to synthesize all evaluation results involving a given training system, and would assist training managers in using the evaluation results to make the desired changes to the system (P. Bruce, personal communication, August 24, 1994).

Finally, development of a comprehensive, integrated *evaluation plan* is needed to meet the diverse informational needs of various curriculum developers, evaluators, training managers, and senior squadron and wing leadership. When implemented, the plan would serve as an effective guide for coordinated information collection, processing, and use (Nullmeyer, McGann, & Rooney, 1986). The specification for such a plan formed part of MATS that was developed by the Laboratory during the 1980s (Fishburne et al., 1987).

Given these problems, we would propose to employ a fully integrated approach to assist the 58 SOW design an I^2MS that could be deployed at the squadron or wing level. As outlined by Nullmeyer, Bruce, and Rockway (1991), four discrete steps would be followed.

First, an I^2MS Working Group would be formed. Such a group would ensure that all users of the information would be active participants in the design process. Working Group members would identify the comprehensive information requirements of the system, and the group would be composed of Air Force and contractor personnel.

Second, a baseline analysis of the existing I^2MS, including both manual and automated components, would be performed. This analysis will provide a common frame of reference to help all participants understand the information requirements, identify potential high-payoff improvements stemming from training interventions, and serve as a benchmark for determining whether improvements in training effectiveness can be linked to the intervention. A variety of data collection methods are required to support this analysis, including over-the-shoulder observations, reviews of archival records, interviews, and surveys (Nullmeyer, Bruce, & Rockway, 1991).

Third, once the baseline analysis has been completed, information requirements for the new I^2MS can be developed. This is done by having the Working Group map the baseline information system onto the desired information elements. By identifying gaps and shortfalls in these elements, the Working Group can then translate existing information requirements into those needed for the new system. Depending on the nature of the shortfalls, requisite changes might include new proficiency data (e.g., new task-based rating forms), new data collection methods (e.g., outside observer judgments), additional analyses (e.g., trends and statistical comparisons), revamped personnel functions (e.g., an Information System manager), automated equipment, or rerouting of reported information. As part of this review process, the Working Group would consider trade-offs in alternative designs for the new I^2MS.

The final step entails supporting on-site development of the new I^2MS, in continuous consultation with squadron and wing users as well as software developers. Depending on the information requirements established in the previous step, I^2MS development might entail some or all of the following: new data collection forms, new data collection methods, software, hardware, and reorganization of functions. Once in place, the I^2MS could be used to assess the effectiveness of any intervention or combination of training-related interventions that the squadron elects to implement.

As a starting point, AFRL/HEA sponsored development of a squadron-based mission information, collection, analysis, and reporting system (MICARS). MICARS is a relational database management system that serves as a living repository of all missions conducted by the host squadron. Key features of MICARS include: a mission-oriented organization that matches the squadron's task event/content requirements; an hierarchical structure that can be interrogated at multiple levels; an internal organization that reflects state-of-the-art theories of human activity systems; an efficient and intuitive user interface; and data collection methods that capture input data at multiple points during mission preparation, execution, and debriefing. MICARS operates in a Filemaker Pro™ software environment and can run on most IBM-compatible desktop or laptop computers. Evaluations of MICARS at KAFB by 15 MH-53J pilots undergoing ART were highly favorable, with a majority of respondents giving the system high marks (i.e., 4.0 or better on a 5-point scale) for interface design, usability, and informativeness (Spiker & Walls, 1996). An operational prototype version of MICARS is presently available at the 58 SOW.

Develop an Advanced Mission Preparation Curriculum

The data from the T^2RM study convincingly show that comprehensive mission planning by the crew is an extremely important factor to ensure mission success. Indeed, the degree of importance placed on mission planning correlates to direct and immediate impact on mission execution performance. These results strongly underscore MP as a very promising area for further research, yet we have only begun to "scratch the surface." For example, does having access to an automated planning system stifle development of critical T^2RM behaviors or does it give planners more time for information-sharing and assumption-testing (Spiker & Nullmeyer, 1995b)? Does access to high fidelity simulation and geospecific imagery overemphasize the technological aspects of MP at the expense of more T^2RM-oriented issues?

Unfortunately, advanced mission planning training is not presently supported at the ART level, as there is a tendency to reduce the combat-mission planning period to a somewhat lesser role in the overall training process. Moreover, there is presently no system within the 58 TRSS to incorporate comprehensive pre-mission planning data, materials, or methods into the mission preparation process. Whether it be rotary-wing (Silverman, 1994), fixed-wing (Spiker et al., in press), or networked (Tourville, et al., in press) training, mission planning is typically viewed as a separate training period from actual simulator training.

For example, in observing MH-53J ART, we noted a tendency for planning activities to be restricted to the "canned" aspects of the mission or to be outright curtailed in some instances. Though some planning was conducted, it was limited to computational tasks, such as weight and balances, verification of time on target, fuel flow, and leg times. The more advanced aspects of planning, such as bump plans, what-if contingencies, and coordination of ground tactics (Spiker, 1995), were rarely addressed (Silverman, 1994).

Accordingly, we recommend development of an enhanced MP curriculum that stresses incorporation of advanced planning materials and semi-automated systems (e.g., SOFPARS) into the mission planning process. As part of this curriculum, we advocate development of methods for students to structure their own "team" coordination activities to promote effective planning activities. Students need to be informed of prescriptive approaches to promote their effective mission team coordination skill, rather than being allowed to operate under a presumption that good performance will "just happen" as they conduct their normal mission planning activities.

As part of ART, we would include an academic review session that covers effective MP procedures and techniques, as well as efficient methods for integrating materials into the planning process. We would extend key T^2RM concepts into the planning phase, where there is ample time and opportunity to identify when and where problem behaviors are being exhibited.

In developing initial course content, we would conduct a detailed analysis of qualitative data items taken from T-MOT during the MP phase of the MC-130P T^2RM study. That study yielded a tremendous amount of useful data where Table 5 presents some exemplary planning behaviors. These have been categorized according to 12 dimensions of MP effectiveness identified during previous research (Spiker & Nullmeyer, 1995a). A complete analysis of that data set--coupled with SME interviews, observations, and literature reviews--would help "populate" the Mission Planning course with T^2RM-related content.

Table 5. Measures of Mission Preparation Effectiveness and Notable T^2RM Behaviors.

Mission Preparation Effectiveness Dimension	Notable T^2RM Behavior from the T-MOT
All planning personnel are effectively utilized	AC asked all crewmembers for "what you need to do your job" and then got it for them
A timeline is established for managing the planning process	AC told crewmembers when they had to be completed with their planning tasks in time for the crew briefing
Precise times are determined for accomplishing the key mission events	Planned AR control time and route backwards from the AR control point. Determined optimal T/O time from these.
High-quality crew briefings are given during various stages of planning	After each crewmember briefs, the AC adds final comments for the crew's consideration
Planning crew achieve an in-depth awareness of threat capabilities along the route	To avoid threats, crew planned to fly very low altitude, terrain mask, and high speed (as necessary) maneuvering
The plan is developed to an appropriate level of detail	FE and CSO prepared the evasion plan of action (Note: a level of detail not provided by many of the crews)
All information sources are checked for recency	AC asked when Intel had last been updated
Information is cross-checked for accuracy and the plan's assumptions are aggressively questioned	AC questions assumptions made in each crewmember's component plan
Ground team and support asset requirements are incorporated into the overall plan	AC modifies plan to incorporate considerations of helicopters for the transload
Mission-essential equipment is well thought out and incorporated into the plan	Crew listed minimum equipment needed to accomplish the mission, such as INS, chaff, flares, etc.
Planning assumptions are subject to extensive "what iffing"	Crew planned to "bump up" their airspeed if they encountered threats during AR
Planners incorporate their real-world experience into the planning process	Crewmembers related their own experiences in the area of operations as they developed the execution plan

Besides training, MP poses some interesting research possibilities for this potentially important area. For example, does having access to an automated planning system stifle development of critical T^2RM behaviors? Or, does such access give planners more time for information-sharing and assumption-testing activities that are so essential for good preparation (Spiker & Nullmeyer, 1995b)? Similarly, does having access to high-fidelity simulation and geo-specific imagery overemphasize the technological aspects of mission planning at the expense of the more personal, T^2RM-oriented issues? These and other questions merit study in their own right, particularly as the SOF community moves closer toward integrating computerized systems into its planning process, both real world and training.

REFERENCES

AFI 36-2243. (June, 1994). AFI-36-2243. *Cockpit/crew resource management program.* Washington DC: Department of the Air Force

Andrews, D.H., Bell, H.H., & Nullmeyer, R.T. (1995). Aircrew team training: A tutorial. In *Proceedings for The 17th Interservice/Industry Training Systems and Education Conference.* Albuquerque, NM.

Bailey, J.G. & Witmer, B.G. (1994). Learning and transfer of spatial knowledge in a virtual environment. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting.* Nashville, TN.

Bruce, P.D. (1989). *Aircrew training evaluation: B-52 and KC-135 formal school training* (AFHRL-TR-88-49, AD A208 860). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Bruce, P.D., Killion, T.H., Rockway, M.R., & Povenmire, H.K. (1991). *B-52 and KC-135 mission qualification and continuation training: A review and analysis* (AL-TR-1991-0010. AD A241 591). Williams AFB, AZ: Armstrong Laboratory, Aircrew Training Research Division.

Department of Defense. (1994, October). *Virtual reality and technologies for combat simulation.* Office of Technology Assessment.

Federal Aviation Administration (1996, May). *Developing and evaluating CRM procedures for a regional air carrier: Phase I report.* (FAA Technical Report). Washington, D.C.: Federal Aviation Administration, Office of the Chief Scientific and Technical Advisor for Human Factors.

Fishburne, Jr., R.P., Williams, K. R., Chatt, J. A., & Spears, W. D. (1987). *Design specification development for the C-130 Model Aircrew Training System: Phase I report (*AFHRL-TR-86-44, AD B115 427). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Gower, D.W., Jr. & Fowlkes, J. (1989). *Simulator sickness in the AH-1S (Cobra) flight simulator.* (USAARL Report No. 89-20). Fort Rucker, AL: Army Aeromedical Research Laboratory.

Harris, R.J. (1994). *ANOVA: An analysis of variance primer.* Itasca, IL: F.E. Peacock Publishers, Inc.

Hays, W.L. (1973). *Statistics.* New York: Holt, Rinehart, and Winston, Inc.

Helmreich, R.L. & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E.L. Weiner, B.G. Kanki, and R.L. Helmreich (Eds.) *Cockpit Resource Management*, San Diego: Academic Press.

Kantowitz, B.H. & Sorkin, R.D. (1983). *Human factors: Understanding people-system relationships*. New York: John Wiley & Sons.

Kennedy, R.S., Lilienthal, M.G., Berbaum, K.S., Baltzley, D.R., & McCauley, M.E. (1989). *Aviation, Space, and Environmental Medicine*, January, 10-16.

Lauber, J.K. (1987). Cockpit resource management: Background studies and rationale. In H.W. Orladay & H.C. Foushee (Eds.), *Cockpit resource management training* (NASA Conference Publication 2455). Moffett Field, CA: NASA-Ames Research Center.

Mirabella, A.,Sticha, P., & Morrison, J. (1997, January). *Assessment of user reactions to the multi service distributed training testbed (MDT2) system.* (Technical Report TBD). Alexandria, VA: US Army Research Institute.

Mowafy, L. & Miller, L. (1993, October). *Visualizing spatial relationships: Training fighter pilots in a virtual environment debrief interface.* (AL/HR-TR-1993-0148, AD A273 972). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Nullmeyer, R. T., Bruce, P. D., & Rockway (1991). Integrated aircrew training management information systems: An organizational perspective. In *Proceedings of the 15th Interservice/ Industry Training Systems and Education Conference.* San Antonio, TX.

Nullmeyer, R. T., McGann, E.J.J., & Rooney, J. (1986). *Test and evaluation strategy: Chronological.* Unpublished document.

Nullmeyer, R. T., Silverman, D. R., Tourville, S. J., & Spiker, V. A. (in press). *Data collection instruments used to measure tactical team resource management (T^2RM) behaviors.* (AFRL-HE-AZ-TR-1998-xxxx) Mesa, AZ: Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Training Research Division.

Nullmeyer, R.T. & Spiker, V.A. (in press). *Utilization strategies for simulation-based mission rehearsal systems* (AL/HR-TR-1996-0147). Mesa, AZ: Armstrong Laboratory, Aircrew Training Research Division.

Nullmeyer, R.T., Bruce, P.D., Conquest, M.T., & Reed, E.T. (1992). The use of the MH-53J weapon systems trainer for mission rehearsal: The mission preparation process and lessons learned. In *Proceedings of the 14th Interservice/Industry Training Systems and Education Conference.* San Antonio, TX.

Pimentel, K. & Teixeira, K. (1993). *Virtual reality: Through the new looking glass*. New York: Intel/Windcrest/MacGraw Hill.

Povenmire, H.K., Rockway, M.R., Bunecke, J.L., & Patton, M.W. (1989). *Evaluation of measurement techniques for aircrew coordination and resource management skills* (UDR-TR-89-108). Williams AFB, AZ: Air Force Human Resources Laboratory, Operations Training Division.

Reed, E.T. (1996). The aerial gunner and scanner simulator: "Affordable virtual reality training for aircrews." In *Proceedings of the Royal Aeronautical Society Conference on Training - Lowering the Cost Maintaining the Fidelity*. London, UK.

Reed, E.T. & Selix, G.R. (1995). High fidelity networked flight simulation for U.S. Air Force special operations and combat rescue training and rehearsal. In *Proceedings of the Royal Aeronautical Society Conference on Flight Simulation Technologies, Capabilities, and Benefits*. London, UK.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw Hill.

Silverman, D. R. (1994). *Combat mission training and simulation: A selected annotated bibliography*. (AL/HR-TR-1994-0172, AD B206 646). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Silverman, D. R. & Spiker, V. A. (1997). A usability assessment of virtual reality simulation for aerial gunner training. *Paper presented at the 41st Annual Meeting of the Human Factors and Ergonomics Society*. Albuquerque, NM.

Silverman, D. R., Spiker, V. A., & Nullmeyer, R. T. (1996). *Human factors evaluation of the Aerial Gunner/Scanner Simulator* (AL/HR-TR-1996-0146). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Silverman, D. R., Spiker, V. A., Tourville, S. J., & Nullmeyer, R. T. (1996). A combat team performance model: Development and initial application. In *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Silverman, D. R., Spiker, V. A., Tourville, S. J., & Nullmeyer, R. T. (1997). Team coordination and performance during combat mission training performance. In *Proceedings of the 19th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Silverman, D. R., Spiker, V. A., Tourville, S. J., & Nullmeyer, R. T. (in press). *Crew position processes effect on team performance during combat mission training* (AL/HR-TR-1997-0136). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V.A. & Nullmeyer, R.T. (1995a). Benefits and limitations of simulation-based mission planning and rehearsal. *Eighth International Symposium on Aviation Psychology*. Columbus, OH.

Spiker, V.A. & Nullmeyer, R.T. (1995b). *Measuring the effectiveness of mission preparation in the Special Operations Forces*. (AL/HR-TR-1995-0071, AD B206 670). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V. A., Nullmeyer, R. T., Tourville, S. J., & Silverman, D. R. (1997). Effects of mission preparation on crew combat mission training performance. *Paper presented at the 41st Annual Meeting of the Human Factors and Ergonomics Society*. Albuquerque, NM.

Spiker, V. A., Silverman, D. R., Tourville, S. J., & Nullmeyer, R. T. (in press). *Tactical team resource management effects on combat mission training performance* (AL/HR-TR-1997-0137). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V. A., Tourville, S. J., & Nullmeyer, R. T. (1997). Network simulation and combat mission training. *Proceedings of the 19th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Spiker, V.A., Tourville, S.J., Silverman, D.R. & Nullmeyer, R.T. (in press). *Candidate AL/HTI research and development issues: A strategic outlook* (AL/HR-TR-1997-xxxx). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V.A., Tourville, S.J., Silverman, D.R., & Nullmeyer, R.T. (1996). *Team performance during combat mission training: A conceptual model and measurement framework* (AL/HR-TR-1996-0092). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Spiker, V.A. & Walls, W.F. (1996, April). *Development of a mission-information collection, analysis, and reporting system (MICARS)* (AL/HR-TR-1996-0043). Santa Barbara, CA: Anacapa Sciences, Inc.

Thurman, R.A. & Mattoon, J.S. (1991). Virtual reality: Theoretical and practical implications. *In Proceedings of the 13th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Tourville, S. J. (1997). Identification of key behaviors for effective team coordination in Special Operations Forces combat mission training. *Paper presented at the Ninth Symposium on Aviation Psychology*. Columbus, OH.

Tourville, S.J., Spiker, V.A., & Nullmeyer, R.T. (1997). *Analysis of the Special Operations Forces network training for joint mission operations simulator training* (AL/HR-TR-1997-0138). Mesa, AZ: Armstrong Laboratory, Human Resources Directorate, Aircrew Training Research Division.

Tourville, S. J., Spiker, V. A., Silverman, D. R., & Nullmeyer, R. T. (1996). An assessment methodology for team combat mission training and rehearsal. In *Proceedings of the 18th Interservice/Industry Training Systems and Education Conference*. Orlando, FL.

Van Cott, H.P. & Warrick, M.J. (1972). Man as a system component. In H.P. Van Cott and R.G. Kinkade (Eds.) *Human engineering guide to equipment design*, 2nd Edition (pp. 17-40). American Institutes for Research: Washington, D.C.

Wickens, C.D. (1984). *Engineering psychology and human performance*. Columbus, OH: Charles Merrill Publishing Company.